Albeverio, S. and Kawasaki, S.
Osaka J. Math.
51 (2014), 1–37

ON A LOCALIZATION PROPERTY OF WAVELET COEFFICIENTS FOR PROCESSES WITH STATIONARY INCREMENTS, AND APPLICATIONS. II. LOCALIZATION WITH RESPECT TO SCALE

SERGIO ALBEVERIO and SHUJI KAWASAKI

(Received August 19, 2011, revised April 18, 2012)

Abstract

Wavelet coefficients of a process have arguments shift and scale. It can thus be viewed as a time series along shift for each scale. We have considered in the previous study general wavelet coefficient domain estimators and revealed a localization property with respect to shift.

In this paper, we formulate the localization property with respect to scale, which is more difficult than that of shift. Two factors that govern the decay rate of cross-scale covariance are indicated. The factors are both functions of vanishing moments and scale-lags. The localization property is then successfully applied to formulate limiting variance in the central limit theorem associated with Hurst index estimation problem of fractional Brownian motion. Especially, we can find the optimal upper bound J of scales $1, \dots, J$ used in the estimation to be $J = 5$ by an evaluation of the diagonal component of the limiting variance, in virtue of the scale localization property.

1. Introduction

Let $Z^n = \{Z_1, Z_2, \dots, Z_n\}$ be a general real-valued stationary ergodic sequence with covariance $r_k = \mathrm{Cov}[Z_{k+1}, Z_1]$, $k \in \mathbb{N}_0$ and finite variance $\sigma^2 = r_0$. Suppose that Z^n is short-range dependent (SRD), i.e. $\sum_{k\in\mathbb{N}} |r_k| < \infty$. Then $\overline{Z}_n = (1/n) \sum_{k=1}^n Z_k$ is a consistent estimator of $\mathbb{E}[Z_1]$ and the limiting variance associated with the central limit theorem (CLT) is given by

$$\lim_{n\to\infty} \mathrm{Var}[\sqrt{n}\,\overline{Z}_n] = \sigma^2 + 2 \sum_{k\in\mathbb{N}} r_k = \sigma^2[1 + 2\delta(\rho)], \tag{1}$$

where $\delta(\rho) = \sum_{k\in\mathbb{N}} \rho_k$ and $\rho = \{\rho_k;\, k \in \mathbb{N}\}$, $\rho_k = r_k/\sigma^2$, is the auto-correlation coefficient of Z^n. Since $\delta(\rho) \equiv 0$ if Z^n is an independent sequence, it turns out that $\delta(\rho)$ represents the adjustment term in order to evaluate $\mathrm{Var}[\sqrt{n}\,\overline{Z}_n]$ correctly when Z^n is not an independent sequence (see Beran (1994), Section 1.1 and Lehmann (1999), Section 2.8).

2000 Mathematics Subject Classification. Primary 60G22, 65T60, 60F05; Secondary 60G18, 60G15, 62F03, 62J10.

Since $\sqrt{1+2\delta(\rho)}$ is the ratio of the confidence interval in estimating $\mathbb{E}[Z_1]$ with respect to that of the case where the sequence Z^n is independent, the confidence interval becomes larger if $\delta(\rho)$ is. This leads to a larger error in hypothesis testing. The evaluation of $\delta(\rho)$ is thus important in general as a measure of the correctness of estimation.

On the other hand, a wavelet-based estimation of parameters associated with stationary-increment processes is formulated in Abry et al. [1]. It is suggested in Tewfik et al. [19] that even though the original process is long-range dependent (LRD), the wavelet coefficient can be made SRD; one can thus always enjoy CLT in wavelet coefficient domain (WCD) estimates, while process domain estimates sometimes result in noncentral limit theorems due to the LRD.

In this paper, we evaluate the limiting variance for CLT in WCD estimates of a parameter, from the point of view of scale localization: A WCD counterpart of (1) is formulated and a localization property—the component corresponding to $\delta(\rho)$ is small—is revealed; The scale localization property is then successfully applied to find the optimal upperbound J of the scales $j = 1, \dots, J$ used in the Hurst index estimates in WCD. This optimization is possible due to the fact that the contribution to the limiting variance in the CLT is almost entirely coming from its diagonal component, so that considering just the diagonal component suffices for the optimization.

So far the localization property of the wavelet coefficient has been considered only for the shift k, essentially (see e.g. Tewfik et al. [19]). Let us call it *k-localization*. The k-localization is obtained easily by the *vanishing moment property* of wavelet, as far as the asymptotic evaluation is concerned. It turns out, however, that a "pointwise" evaluation as given in Albeverio et al. [2] gives rise to a true power of WCD estimates. In fact, the pointwise evaluation is applied to show that the simple WCD estimator by the moment method has a variance which is nearly as small as for the maximum likelihood estimator.

In this paper, we evaluate the j-localization and do it "pointwise". The j-localization is not straightforward like k-localization and is rather difficult. However, by this "pointwise" evaluation, we obtain that the optimal J is 5, which is enabled by the true power of WCD estimate due to the j-localization.

The organization of the paper is as follows. Section 2 is the preliminary, in which we give concepts related to WC, its covariance, parameter estimates in WCD and reformulation of the CLT associated with the estimates. Section 3 presents one of the main results that states the basic j-localization theorem for the covariance itself, evaluating its decay as a function of scale-lags, shift-lags and vanishing moments. Section 4 explains what is to be proved essentially and what the difficulty is beyond the k-localization. Section 5 gives evaluations of two key elements appearing in the basic j-localization. In Section 6, as an application of the basic j-localization theorem in Section 3, we evaluate the limiting variance in the CLT of the Hurst index estimation for FBM. This evaluation may be considered as a functional form of the j-localization. Section 7 is devoted to the proof of the basic j-localization theorem, Theorem 2. Section 8 presents the proofs of Theorem 3 which is the fundamental idea of the j-localization, and of Theorem 7 which determines the optimal scale upper bound in the Hurst index estimation stated above. Sections 9 and 10 are the proofs of propositions and lemmas, respectively. The last Section 11 is the concluding remarks.

2. Preliminaries

We present some basics of the paper in the following subsections.

We have considered WCD estimates of parameters of stationary-increment process in [2]. In this subsequent study we assume that $X = \{X_t \mid t \geq 0\}$ is a process with H-self similarity (H-ss) and Gaussianity as well as with stationary-increments (si), i.e. X is a fractional Brownian motion (FBM) with Hurst index $0 < H < 1$. We will mention a reason of this restriction on the process later. Also, let $X^T = \{X_t \mid 0 \leq t \leq T\}$ be an observed sample path of X. Thus we consider the wavelet coefficients of this X or X^T.

2.1. Wavelet coefficients and assumptions on wavelet. Let ψ be a real-valued wavelet on $\mathbb{R}_+$ satisfying the assumption that it has
($\psi 1$) compact support on $W = [0, w]$ for some real $w \geq 1$, and is bounded;
($\psi 2$) γ_0-th order vanishing moment for some $\gamma_0 \in \mathbb{N}$: $\int_{\mathbb{R}_+} t^r \psi(t)\, dt = 0$, $r = 0, 1, \ldots, \gamma_0 - 1$.

Let $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$. The variables j and k are integers and called *scale* and *shift*, respectively. The support of $\psi_{j,k}(t)$ is $\mathrm{supp}(\psi_{j,k}) = [2^j k, 2^j(k+w)]$. Let $\mathcal{J} = \{J_0+1, \ldots, J\}$, $J_0 < J$ denote the range of those j that are used in the estimator. In fact we consider those WCD estimators that are of the form given by (9) below. Although examples of estimators with $|\mathcal{J}| = \infty$ are given in Albeverio et al. [2] as well, we consider only the case of finite $\mathcal{J}$ in this paper. This is due to the $|\mathcal{J}|$-dimensional CLT for WCD estimators discussed below. Nevertheless, the study of the case of finite $\mathcal{J}$ is useful. This is because, in case of infinite $\mathcal{J}$, it may often be truncated to finite range with certain prescribed error, at least practically. On the other hand, the range of k is restricted by the assumption ($\psi 1$) on the initial k. We will mention this later.

General theories of wavelet analysis of stochastic processes can be found in Tanaka [17] [18], Chiann [7], Xie et al. [20] for example.

In applications, the way of notation of vanishing moment as in ($\psi 2$) (r ranging up to $\gamma_0 - 1$, not γ_0) is often used so far. This is because, with this notation, the covariance decay of the wavelet coefficients of processes with H-ss is given as $O(k^{2H-2\gamma_0})$ as $k \to \infty$ (k: lag of shifts), in which an argument that the terms $((s-t)/k)^r$, $r = 0, \ldots, \gamma_0 - 1$ vanish, is involved. However, since several essential quantities in this paper are written not in terms of γ_0 but $\gamma_0 - 1$, we will use the notation $\gamma \triangleq \gamma_0 - 1$ as well.

For each $T > 0$, let $N_{j,T} \in \mathbb{N}$ be the number of wavelet coefficients that are available up to $t = T > 0$, for each j: $N_{j,T} = \max\{k \mid 2^j(w+k) \leq T\} = \lfloor 2^{-j}T - w \rfloor$. We denote by $S^T = \{s_j^T \mid j \in \mathcal{J}\}$ the random vectors of wavelet coefficients $s_j^T = \{s_j(k) \mid k = 1, \dots, N_{j,T}\}$, defined by

$$s_j(k) = \int_{2^j k}^{2^j(k+w)} \psi_{j,k}(t) X(t)\, dt.$$

Although a process may be observed in discrete time by engineering sensor systems from realistic point of view, we consider here, as a first foundation of the relevant theory, in continuous time setting and thus the wavelet coefficient as above. Wavelet coefficient of this form is considered by many authors (see, e.g. [1] [15] [19] [20]). This is because we can use analytical methods like Fourier transform techniques and differentiation and integration to avoid nonessential difficulties in calculation itself. Rather, it may be important to establish the scenario of localization with respect to scale in the first stage. Such scope of this paper indeed reveals what happens in the scenario, thus providing a significant value to the localization theorem, beyond the assumption of observation being either discrete or continuous time. Also, it is often the case that a statistical analysis of stochastic processes is first developed either in discrete or continuous time and then a similar structure is found in the other.

Now, let us recall the correspondence of wavelet parameters and process properties (see e.g. Flandrin [10]). For shift k and si, we have that the time series s_j^T is, for each j stationary: For $\nu \in \mathbb{N}$, $n \in \mathbb{N}$ and $\{k_1, \dots, k_n\} \subset \mathbb{N}_0$,

$$(s_0(k_1 + \nu), \dots, s_0(k_n + \nu)) \overset{(d)}{=} (s_0(k_1), \dots, s_0(k_n)),$$

where $\overset{(d)}{=}$ stands for equality in distribution. The stationarity of s_j^T is used in the CLT for WCD estimates. So the assumption of X having si is essential.

On the other hand, we have used the assumption of H-ss and Gaussianity of X just for the sake of simplicity of formulation. In fact, as for scale j and H-ss, we have, for $j \in \mathbb{Z}$, $n \in \mathbb{N}$ and $\{k_1, \dots, k_n\} \subset \mathbb{N}_0$,

$$(s_j(k_1), \dots, s_j(k_n)) \overset{(d)}{=} 2^{(H+1/2)j}(s_0(k_1), \dots, s_0(k_n)).$$

By this, it turns out that one can consider only s_0^T whenever a single j is concerned. As for Gaussianity, it is used just for an explicit evaluation of the covariance, $\mathrm{Cov}[H_l(X), H_l(Y)] = (\mathrm{Cov}[X, Y])^l$, $l \in \mathbb{N}$, for a 2-dimensional Gaussian r.v. (X, Y) and an l-th order Hermite polynomial $H_l(x)$. Similar arguments without the two assumptions might be possible, but even though we indeed put the assumptions, the theoretical development here is not easy, and is even more complicated if we do not make these assumptions. Therefore, in this paper we describe our results not in the most general but in a simple way, with the additional two assumptions, in order to make the essence of our argument clear. Finally, we remark that if X has mean 0, then so is S.

2.2. Covariance of wavelet coefficients. Let $j, j' \in \mathcal{J}$ and let us write $j' = j + m$, $m = 0, 1, \ldots, J - j$. This m corresponds to the lag of scales. By the H-ss of X, we have

$$\begin{aligned}
&(s_j(k_1), s_{j'}(k_2)) \\
&= \left(2^{j/2} \int_W \psi(s) B_H(2^j(s+k_1))\, ds,\ 2^{j'/2} \int_W \psi(t) B_H(2^{j'}(t+k_2))\, dt\right) \\
&\overset{(\mathrm{d})}{=} \left(2^{(H+1/2)j} \int_W \psi(s) B_H(s+k_1)\, ds,\ 2^{(H+1/2)j+m/2} \int_W \psi(t) B_H(2^m(t+k_2))\, dt\right).
\end{aligned}$$

We recall that the covariance of the wavelet coefficients of X can then be written as (see, e.g. Flandrin [10])

$$\begin{aligned}
&\mathrm{Cov}[s_j(k_1), s_{j'}(k_2)] \\
&= 2^{(2H+1)j+(m/2)} \left(-\frac{1}{2}\right) \iint_{W^2} \psi(s)\psi(t) |s - 2^m t + k_1 - 2^m k_2|^{2H}\, ds\, dt \\
&= 2^{(2H+1)j} \cdot \mathrm{Cov}[s_0(k_1), s_m(k_2)] \\
&\triangleq 2^{(2H+1)j} r(m, n),
\end{aligned}$$

for $n = k_1 - 2^m k_2 \in \mathbb{Z}$. Especially, if $j = j'$, then

$$\mathrm{Cov}[s_j(k_1), s_j(k_2)] = 2^{(2H+1)j} \cdot \mathrm{Cov}[s_0(k_1), s_0(k_2)];$$

further, if $k_1 = k_2$, then $\sigma_j^2 \triangleq \mathrm{Var}[s_j(1)] = 2^{(2H+1)j} \cdot \mathrm{Var}[s_0(1)] = 2^{(2H+1)j}\sigma_0^2$. We write $r(0, n) = r(n)$.

Let $\mathsf{s}_j(k)$ be the normalized wavelet coefficient: $\mathsf{s}_j(k) = s_j(k)/\sigma_j$. Then, we have

$$\begin{aligned}
\mathrm{Cov}[\mathsf{s}_j(k_1), \mathsf{s}_{j'}(k_2)]_{k_1 - 2^m k_2 = n} &= \mathrm{Cov}[\mathsf{s}_0(k_1), \mathsf{s}_m(k_2)] \\
&\triangleq \rho(m, n) \\
&= \frac{2^{-Hm}}{\sigma_0^2} \left(-\frac{1}{2}\right) \iint_{W^2} \psi(s)\psi(t) |s - 2^m t + n|^{2H}\, ds\, dt \qquad (2) \\
&= 2^{-(H+1/2)m} \frac{r(m, n)}{\sigma_0^2}.
\end{aligned}$$

This $\rho(m, n)$ is the correlation coefficient of $s_j(k_1)$ and $s_{j'}(k_2)$, with indices m and n corresponding to lags of scales and shifts, respectively. We set $\rho(n) \triangleq \rho(0, n)$ especially.

Here we remark that there is a restriction on the shift index k, according to the assumption (ψ1). We recall that a mother wavelet function ψ associated with a multi-resolution analysis (MRA) is generated by a filter

$$m_0(\lambda) = \sum_{k=N_1}^{N_2} h_k e^{-ik\lambda}, \quad \lambda \in \mathbb{R} \tag{3}$$

for some $N_1, N_2 \in \mathbb{Z}$ and $\{h_k\} \subset \mathbb{R}$, and has its support on $[(N_1 - N_2 + 1)/2, (N_2 - N_1 + 1)/2]$ (Härdle et al. [12]). In order to take the wavelet coefficients $s_j(k)$ of X^T, we consider only those $k = k_0, k_0+1, \dots$ that satisfy $\operatorname{supp}(\psi_{j,k_0}) \subset [0, T]$. Here we have

$$\operatorname{supp}(\psi_{j,k_0}) = \left[2^j\left(\frac{N_1 - N_2 + 1}{2} + k_0\right), 2^j\left(\frac{N_2 - N_1 + 1}{2} + k_0\right)\right]. \tag{4}$$

Thus the initial shift k_0 must be

$$k_0 = \left\lceil \frac{N_2 - N_1 - 1}{2} \right\rceil, \tag{5}$$

where $\lceil x \rceil$ is the least integer larger than or equal to $x \in \mathbb{R}$. Especially, in the case of a Daubechies wavelet, we have $N_2 - N_1 = 2\gamma - 1$ (Härdle et al. [12]) so that $k_0 = \gamma - 1$. For such k_0, we take the shift $k = k_0, k_0 + 1, \dots$. We will see later that this restriction on k_0 is essential in the j-localization theorem.

2.3. WCD estimates and CLT associated with it. The general WCD estimator we consider in this paper is as follows (Albeverio et al. [2]). Suppose that we want to estimate a parameter $\zeta \in \mathbb{R}$ associated with X, which can be written in the form $\zeta = f(\boldsymbol{\theta})$ for a given $f \colon \mathbb{R}^{\mathcal{J}} \mapsto \mathbb{R}$, with $\boldsymbol{\theta} = (\theta_j)_{j \in \mathcal{J}} \in \mathbb{R}^{\mathcal{J}}$. Here θ_j, $j \in \mathcal{J}$ are assumed to be a functional expectation of the wavelet coefficient at scale j, $\theta_j = \mathbb{E}[g(s_j(1))]$ for some $g \colon \mathbb{R} \mapsto \mathbb{R}$ with Hermite expansion $g(x) = \sum_{l \geq p} c_l H_l(x)$ in $L^2(\mathbb{R}, e^{-x^2/2} dx/\sqrt{2\pi})$ for some Hermite rank $p \geq 1$.

As an estimator $\hat{\boldsymbol{\theta}}_T = (\hat{\theta}_{j,T})_{j \in \mathcal{J}} \in \mathbb{R}^{\mathcal{J}}$ of $\boldsymbol{\theta}$, we especially take the one defined by

$$\hat{\theta}_{j,T} = \frac{1}{N_{J,T}} \sum_{n=1}^{N_{J,T}} Y_j(n), \quad Y_j(n) = \frac{1}{d_j} \sum_{k=1}^{d_j} g(s_j(d_j n + k)), \tag{6}$$

where $d_j = 2^{J-j}$. Since s_j^T is stationary and ergodic, so is $\boldsymbol{Y}(n) = (Y_j(n))_{j \in \mathcal{J}}$, and hence $\hat{\boldsymbol{\theta}}_T$ is a consistent and unbiased estimator of $\boldsymbol{\theta}$: $\hat{\boldsymbol{\theta}}_T \to \boldsymbol{\theta}$ a.s. as $T \to \infty$ and $\mathbb{E}[\hat{\boldsymbol{\theta}}_T] = \boldsymbol{\theta}$. Then we assume that ζ is estimated consistently by $\hat{\zeta}_T = f(\hat{\boldsymbol{\theta}}_T)$:

$$f(\hat{\boldsymbol{\theta}}_T) \to f(\boldsymbol{\theta}) \quad \text{a.s.} \quad \text{as } T \to \infty. \tag{7}$$

Thus the WCD estimator $\hat{\zeta}_T$ we consider here is of the form $f(\{g(s_j^T)\}) = f \circ g(s^T)$. Examples of WCD estimators of this form are given in Albeverio et al. [2].

As a consistent estimator $\hat{\boldsymbol{\theta}}_T$, several authors have considered the one of the form $\left\{N_{j,T}^{-1}\sum_{k=1}^{N_{j,T}} g(s_j(k))\right\}_{j\in\mathcal{J}}$ so far (see e.g. [1]). An inconvenience of this estimator is that $(s_j(k))_{j\in\mathcal{J}}$, $k = 1, 2, \ldots$ is not a stationary vector with respect to k. The reason why we take the estimator of the renormalized form as in (6) is so as to make $\boldsymbol{Y}(n)$ a stationary vector.

The CLT for WCD estimates was first considered in case of Hurst index estimation by Bardet et al. [3] and then in a general case by Albeverio et al. [2]. To study the localization property through WCD counterpart of (1), we begin our argument here by reformulating the limiting variance in the CLT associated with (7). This makes it apparent that the elements of the limiting covariance matrix $\boldsymbol{\Sigma}$ do not depend on j but only on the scale-lags m.

In Albeverio et al. [2, Theorem 3], essentially a CLT of the following form for the renormalized sequence $Y_j(n)$ is considered: If $(\gamma_0 - H)p > 1$, then

$$\sqrt{N_{J,T}}[\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}] \Rightarrow \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \text{as} \quad T \to \infty, \tag{8}$$

with $\hat{\boldsymbol{\theta}}_T = (\hat{\theta}_{j,T})_{j\in\mathcal{J}}$ and $\hat{\theta}_{j,T} = N_{J,T}^{-1}\sum_{n=1}^{N_{J,T}} Y_j(n)$. Here $\gamma_0 \geq 2$ is sufficient for $(\gamma_0 - H)p > 1$, for all $0 < H < 1$ and $p \geq 1$. For example, $(\gamma_0, p) = (1, 1)$ cannot be a sufficient condition.

We set $\hat{\boldsymbol{\theta}}'_T = (2^{-j/2}\hat{\theta}_{j,T})_{j\in\mathcal{J}}$ and $\boldsymbol{\theta}' = (2^{-j/2}\theta_j)_{j\in\mathcal{J}}$ instead of $\hat{\boldsymbol{\theta}}_T$ and $\boldsymbol{\theta}$. As the assumption on the class of functions $f\colon \mathbb{R}^{\mathcal{J}} \mapsto \mathbb{R}$, we especially consider those of the linear form

$$f(\boldsymbol{\theta}) = \sum_{j\in\mathcal{J}} a_j\varphi_j(\theta_j), \quad \{a_j\} \subset \mathbb{R}, \quad \varphi_j : \mathbb{R} \mapsto \mathbb{R}, \tag{9}$$

in this paper. All the examples of WCD estimators in Albeverio et al. [2] satisfy this assumption. For this form of the estimator, we can see how the effect of l-th Hermite polynomials is involved in the quantities considered in this paper, through the following argument. Let $\dot{\varphi}_j(x) = d\varphi_j(x)/dx$. We can write

$$f(\hat{\boldsymbol{\theta}}_T) - f(\boldsymbol{\theta}) = \sum_{j\in\mathcal{J}} a_j \frac{\varphi_j(\hat{\theta}_{j,T}) - \varphi_j(\theta_j)}{\hat{\theta}_{j,T} - \theta_j} \cdot (\hat{\theta}_{j,T} - \theta_j) \simeq \sum_{j\in\mathcal{J}} a_j[\theta_j\dot{\varphi}_j(\theta_j)] \cdot \frac{\hat{\theta}_{j,T} - \theta_j}{\theta_j},$$

when T is large (where $\simeq$ stands for equality modulo $o(T)$). Here, if θ_j are moments of even order, $\theta_j = \mathbb{E}[s_j^{2\nu}(1)]$ for some $\nu \in \mathbb{N}$, we can write

$$\frac{\hat{\theta}_{j,T} - \theta_j}{\theta_j} = \frac{1}{N_{j,T}}\sum_{k=1}^{N_{j,T}}\sum_{l\in\mathbb{N}} c_l H_l(s_j(k))$$

for some $\{c_l\}$ in general. For the sake of simplicity, let us consider $g(x) = H_l(x)$ for a fixed positive integer l. The case of general $g(x)$ can be treated by summation of such terms with respect to l. By this we can avoid nonessential complexities. Because of this, we sometimes suppress the subscript l from the relevant symbols appearing below.

Now we state the general CLT of WCD estimates, which is a reformulation of the one given in Albeverio et al. [2].

Proposition 1 (Reformulation from [2]). *Let $g(x) = H_l(x)$ for a fixed l. If $(\gamma_0 - H)p > 1$, then the $|\mathcal{J}|$-dimensional CLT*

$$\sqrt{T}[\hat{\boldsymbol{\theta}}'_T - \boldsymbol{\theta}'] \Rightarrow \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \tag{10}$$

and the one-dimensional CLT

$$\sqrt{T}[f(\hat{\boldsymbol{\theta}}'_T) - f(\boldsymbol{\theta}')] \Rightarrow \mathcal{N}(0, \upsilon^2_{H,\mathcal{J}}) \tag{11}$$

hold respectively. The limiting covariance matrix $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{j,j'})$ and the limiting variance $\upsilon^2_{H,\mathcal{J}}$ are given by

$$\boldsymbol{\Sigma}_{j,j+m} = 2^{-m/2} R_H(m) \quad \text{with} \quad R_H(m) = \rho^l(m,0) + 2\sum_{k\in\mathbb{N}} \rho^l(m,k)$$

for $m = 0, 1, \ldots$ and

$$\upsilon^2_{H,\mathcal{J}} = \xi_{\mathcal{J}} R_H + 2\sum_{m=1}^{|\mathcal{J}|-1} \xi_{\mathcal{J}}(m) R_H(m) \tag{12}$$

with

$$\xi_{\mathcal{J}}(m) \triangleq \sum_{j=J_0+1}^{J-m} 2^j a_j a_{j+m} [\theta_j \dot{\varphi}_j(\theta_j)][\theta_{j+m}\dot{\varphi}_j(\theta_{j+m})], \tag{13}$$

respectively. Here we have set $\xi_{\mathcal{J}} \triangleq \xi_{\mathcal{J}}(0) = \sum_{j\in\mathcal{J}} 2^j a_j^2 [\theta_j \dot{\varphi}_j(\theta_j)]^2$ and $R_H \triangleq R_H(0) = 1 + 2\sum_{k\in\mathbb{N}} \rho^l(k)$, respectively.

We remark that since $N_{j,T} \simeq 2^{-j}T$ and $N_{J,T} d_j \simeq N_{j,T}$ for large T, we have

$$\hat{\theta}'_{j,T} = \frac{2^{-j/2}}{N_{J,T}} \sum_{n=1}^{N_{J,T}} \left[\frac{1}{d_j} \sum_{k=1}^{d_j} g(s_j(d_j n + k)) \right] \simeq \frac{2^{-j/2}}{N_{j,T}} \sum_{k=1}^{N_{j,T}} g(s_j(k)) \triangleq \hat{\theta}^*_{j,T}$$

or, more precisely, $Pr\left(\left\{\left|\hat{\theta}'_{j,T} - \hat{\theta}^*_{j,T}\right| > \varepsilon\right\}\right) \to 0$ as $T \to \infty$. The convergence in probability in turn implies $Pr(\{|{}^t\boldsymbol{x}\cdot\hat{\boldsymbol{\theta}}'_T - {}^t\boldsymbol{x}\cdot\hat{\boldsymbol{\theta}}^*_T| > 0\}) \to 0$ as $T \to \infty$, for all ${}^t\boldsymbol{x} = (x_{J_0+1}, \ldots, x_J) \in \mathbb{R}^{\mathcal{J}}$. Hence we can modify (10) further as

$$\sqrt{T}[\hat{\boldsymbol{\theta}}^*_T - \boldsymbol{\theta}'] \Rightarrow \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

by the Cramer–Wald device in Billingsley [5] and [6, Theorem 3.1]. Also, $\sqrt{T}[\hat{\boldsymbol{\theta}}^*_T - \boldsymbol{\theta}']$ has asymptotically the same distribution as

$$\{\sqrt{N_{j,T}}\,[\hat{\theta}^*_{j,T} - \theta_j]\}_{j\in\mathcal{J}} = \left\{\frac{1}{\sqrt{N_{j,T}}}\sum_{k=1}^{N_{j,T}}[g(s_j(k)) - \theta_j]\right\}_{j\in\mathcal{J}}. \tag{14}$$

While $\hat{\theta}_{j,T}$ has the merit of a theoretical formulation in the sense that $\boldsymbol{Y}(n)$ is a stationary vector, $\hat{\theta}^*_{j,T}$ may be more convenient for practical calculations.

In (12), we have

$$v^2_{H,\mathcal{J}} = \xi_{\mathcal{J}} R_H[1 + 2\delta_\xi(\rho_R)] \triangleq C_{H,\mathcal{J}}\xi_{\mathcal{J}} R_H, \tag{15}$$

where

$$\delta_\xi(\rho_R) = \sum_{m=1}^{|\mathcal{J}|-1}\frac{\xi_{\mathcal{J}}(m)}{\xi_{\mathcal{J}}}\cdot\frac{R_H(m)}{R_H} = \sum_{m=1}^{|\mathcal{J}|-1} 2^{-(H+\frac{1}{2})ml}\cdot\frac{\xi_{\mathcal{J}}(m)}{\xi_{\mathcal{J}}}\cdot\frac{\bar{R}_H(m)}{\bar{R}_H},$$

with $\bar{R}_H(m) \triangleq r^l(m,0) + 2\sum_{k\in\mathbb{N}} r^l(m,n)$ and $\bar{R}_H = \bar{R}_H(0)$. We will show below that this $\delta_\xi(\rho_R)$ is indeed small, which is part of the j*-localization theorem*.

The relation (15) may be considered to be a WCD counterpart of (1). To be more precise, let $\dot{\boldsymbol{f}} = (\dot{\varphi}_{J_0+1},\dots,\dot{\varphi}_J)$ and $\boldsymbol{\Lambda}$ be a $|\mathcal{J}|\times|\mathcal{J}|$-matrix,

$$\boldsymbol{\Lambda} = \operatorname{diag}(\boldsymbol{\Sigma}) \equiv R_H \quad \text{if} \quad j = j'; \quad = 0 \quad \text{otherwise}. \tag{16}$$

Then if $\delta_\xi(\rho_R)$ is indeed small, we may have

$$\underline{C}_{H,\mathcal{J}}\,{}^t\dot{\boldsymbol{f}}\boldsymbol{\Lambda}\dot{\boldsymbol{f}} \le v^2_{H,\mathcal{J}} = {}^t\dot{\boldsymbol{f}}\boldsymbol{\Sigma}\dot{\boldsymbol{f}} \le \overline{C}_{H,\mathcal{J}}\,{}^t\dot{\boldsymbol{f}}\boldsymbol{\Lambda}\dot{\boldsymbol{f}} \tag{17}$$

for some $0 < \underline{C}_{H,\mathcal{J}} < 1 < \overline{C}_{H,\mathcal{J}}$ with $\inf\overline{C}_{H,\mathcal{J}} = C_{H,\mathcal{J}}$. This is the expression of j-localization in the case of limiting variance of CLT in the WCD estimate. We remark that ${}^t\dot{\boldsymbol{f}}\boldsymbol{\Lambda}\dot{\boldsymbol{f}} = R_H\|\dot{\boldsymbol{f}}\|^2 = \xi_{\mathcal{J}} R_H$. When the cross-scale correlation sums $R_H(m)$, $m = 1,2,\dots$ are small, $v^2_{H,\mathcal{J}}$ is almost the same as the auto-scale correlation sum $\xi_{\mathcal{J}} R_H$. We are especially interested in cases in which the bounds are sufficiently tight with $\underline{C}_{H,\mathcal{J}}$ and $\overline{C}_{H,\mathcal{J}}$ close to 1. In such cases, the covariance matrix $\boldsymbol{\Sigma}$ is "close" to the diagonal matrix $\boldsymbol{\Lambda}$, i.e. the entries $\hat{\theta}_{j,T}$ of $\hat{\boldsymbol{\theta}}_T$ are "close" to be independent.

If the k-localization is desired as well as j-localization, which may be considered a *time-frequency* simultaneous localization, one may proceed further on the j-localization expression (17). In (16), $R_H \equiv \lim_{T\to\infty}\operatorname{Var}\big[\sqrt{N_{j,T}}\hat{\theta}^*_{j,T}\big] = \lim_{T\to\infty}{}^t\boldsymbol{u}_{N_{j,T}}\boldsymbol{\Gamma}_{N_{j,T}}\boldsymbol{u}_{N_{j,T}}$, where $\boldsymbol{u}_N = (1/\sqrt{N},\dots,1/\sqrt{N})$ (N terms) and

$$\boldsymbol{\Gamma}_N = ((\operatorname{Cov}[s_0(k_1), s_0(k_2)])^l)_{1\le k_1,k_2\le N} \equiv ((\operatorname{Cov}[s_0(|k_1-k_2|+1), s_0(1)])^l)_{1\le k_1,k_2\le N}.$$

This $\mathbf{\Gamma}_{N_{j,T}}$ satisfy the k-localization, as shown in [2], so that $c_{*H}\boldsymbol{I}_N \le \mathbf{\Gamma}_{N_{j,T}} \le c^*_H \boldsymbol{I}_N$ for the $N \times N$ identity matrix $\boldsymbol{I}_N$ (The identity matrix stems from 1 in the definition of R_H). Thus (17) can further be written as

$$C^1_{H,\mathcal{J}}\xi_{\mathcal{J}} \le v^2_{H,\mathcal{J}} \le C^2_{H,\mathcal{J}}\xi_{\mathcal{J}}, \tag{18}$$

where $C^1_{H,\mathcal{J}} = c_{*H}\underline{C}_{H,\mathcal{J}}$ and $C^2_{H,\mathcal{J}} = c^*_H\overline{C}_{H,\mathcal{J}}$, respectively. Since $\xi_{\mathcal{J}}$ represents the variance of least square estimates in WCD, (18) is a reasonable one. Therefore, after obtaining the k- or j-localization bounds, $\xi_{\mathcal{J}}$ should be evaluated sufficiently precisely.

It is important to note that the term $2\delta_\xi(\rho_R)$ in (15) is small is by far a stronger assertion than the known asymptotic decay of the single covariance term (see e.g. Tewfik et al. [19])

$$2^{-m/2}r(m,n) = O(n^{-2(\gamma_0-H)}), \quad \text{as} \quad n = |2^m k_1 - k_2| \to \infty.$$

In the application in Section 6 below, which indeed need that $\delta_\xi(\rho_R)$ is small, the asymptotic decay does not work. Thus by *localization*, we mean that the summation defining $\delta_\xi(\rho_R)$, which is over non-diagonal cross-scale covariance components, is small, not just that the asymptotic decay of the covariance is fast. Although we will prove only the upper bound of (17), the result will turn out to be enough for the scope here. We will write $C_{H,\mathcal{J}}$ for $\overline{C}_{H,\mathcal{J}}$ hereafter and show that $\delta_\xi(\rho_R)$ is small so that $C_{H,\mathcal{J}}$ is close to 1.

The ratio $\bar{R}_H(m)/\bar{R}_H$ is common to all processes with H-ss and si, whereas $\xi_{\mathcal{J}}(m)/\xi_{\mathcal{J}}$ depends on each estimation problem. In the present case, the estimation problem is a linear least-square regression. We will consider the evaluation of the two ratios $\bar{R}_H(m)/\bar{R}_H$ and $\xi_{\mathcal{J}}(m)/\xi_{\mathcal{J}}$ in the rest of the paper, to give the evaluation of $v^2_{H,\mathcal{J}}$ through the "diagonal" component $\xi_{\mathcal{J}}R_H$. We may consider it as a functional form of j-localization. We first establish, in the next theorem, its basic form for correlation coefficients itself.

3. Main result (1)—localization of wavelet coefficient with respect to the scale j

We recall that a wavelet ψ that is generated by the two-scale relation in MRA is given by (Daubechies [8])

$$\hat{\psi}(\lambda) = e^{-i\lambda/2}\overline{m_0\left(\frac{\lambda}{2}+\pi\right)}\hat{\phi}\left(\frac{\lambda}{2}\right), \tag{19}$$

where m_0 is given by (3) and $\hat{\phi}(\lambda) = \prod_{j=1}^{\infty} m_0(2^{-j}\lambda)$ in $L^2(\mathbb{R})$.

Theorem 2. *Let ψ be associated with an MRA. For $0 < H < 1$, $m = 1, 2$ and $\gamma \in \mathbb{N}$, we have*

$$2^{-m/2}\frac{r(m,n)}{\sigma_0^2} \le \Psi_{\gamma,H}(m) K_\gamma(m,n), \quad n \in \mathbb{N}_0, \tag{20}$$

where $K_\gamma(m,n)$ and $\Psi_{\gamma,H}$ are given as follows:
(i) *For $m = 1$,*

$$K_\gamma^2(1,n) \triangleq \sum_{k\in\mathbb{N}} \{1 + (n + 2^m k)^2\}^{-2\gamma}\Big|_{m=1}, \tag{21}$$

while

$$\Psi_{\gamma,H}(1) \triangleq A_\gamma(1) + 2^{2H}\left\{\frac{5}{(1+2\gamma)^2}\right\}^\gamma B_\gamma(1)$$

where $A_\gamma(m)|_{m=1}$ and $B_\gamma(m)|_{m=1}$ are such that

$$A_\gamma^2(1) = \frac{1}{\pi}\int_{(\pi,2\pi]}\left|\frac{\hat\psi(2\lambda)}{\hat\psi(\lambda)}\right|^2 d\lambda \quad \text{and} \quad B_\gamma^2(1) = \frac{1}{\pi}\int_{(0,\pi]}\left|\frac{\hat\psi(\lambda)}{\hat\psi(2\lambda)}\right|^2 d\lambda,$$

respectively.
(ii) *For $m = 2$,*

$$K_\gamma^2(2,n) \triangleq \sum_{k_1\in\mathbb{N}}\sum_{k_2\in\mathbb{N}} \{1 + (n + 2^m(k_1+k_2))^2\}^{-2\gamma}\Big|_{m=2}, \tag{22}$$

while

$$\begin{aligned}\Psi_{\gamma,H}(2) \triangleq \frac{1}{\sqrt{32(2\gamma-1)}}\Bigg[&4^{2H}\left\{\frac{17}{1+(6\gamma+1)^2}\right\}^{(2\gamma-1)/2} \cdot B_\gamma(1)B_\gamma(2,1)\\ &+ 2^{2H}\left\{\frac{17}{1+(2\gamma+3)^2}\right\}^{(2\gamma-1)/2} \cdot A_\gamma(1)B_\gamma(2,2) + A_\gamma(1)A_\gamma(2)\Bigg],\end{aligned}$$

where $B_\gamma(m,\nu)|_{m=2}$, $\nu = 1, 2$ and $A_\gamma(2)$ are such that

$$[B_\gamma(2,\nu)]^2 \triangleq \frac{1}{\pi/2}\int_{((\nu-1)\pi/2,\nu\pi/2]}\left|\frac{\hat\psi(\lambda)}{\hat\psi(2\lambda)}\right|^2 d\lambda, \quad \nu = 1, 2,$$
$$[A_\gamma(2)]^2 \triangleq \frac{1}{\pi/2}\int_{(\pi,3\pi/2]}\left|\frac{\hat\psi(2\lambda)}{\hat\psi(\lambda)}\right|^2 d\lambda.$$

The factors $\Psi_{\gamma,H}(m)$ and $K_\gamma(m,n)$ that determine the decay rate of the covariance $r(m,n)$ with respect to scale m, shift n and the vanishing moment γ, are related to MRA wavelet functions and stationarity of increments of X^T, respectively.

The factor $\Psi_{\gamma,H}$ consists of several terms. For $m = 1$, the second term has an "overhead" 2^{2H}, but it is reduced by a "compensation" term $\{5/(1+2\gamma)^2\}^{\gamma}$. This is similar for $m = 2$ as well.

We consider (20) only for $m = 1$ and 2. This is because these values of m are sufficient for the scope of the present paper, while the general formulation of the evaluation for $m \geq 3$ is not easy. For a fixed $m \geq 3$, the evaluation may be possible but quite involved, due to the fact that the functions $\psi(2\lambda)/\psi(\lambda)$, $\psi(2^2\lambda)/\psi(2\lambda), \ldots, \psi(2^m\lambda)/\psi(2^{m-1}\lambda)$ or $\psi(\lambda)/\psi(2\lambda)$, $\psi(2\lambda)/\psi(2^2\lambda), \ldots, \psi(2^{m-1}\lambda)/\psi(2^m\lambda)$ shrink on the λ-axis towards $\lambda = 0$ in different manners.

Let $K_\gamma(m)$, $m = 1, 2$ be defined by $K_\gamma^l(m) \triangleq K_\gamma^l(m,0) + 2\sum_{n\in\mathbb{N}} K_\gamma^l(m,n)$.

Theorem 3. *With the same setting as in* Theorem 2, *we have*

$$R_H(m) \leq \frac{1}{C_*^{(1)}} \cdot [\Psi_{\gamma,H}(m) K_\gamma(m)]^l \cdot R_H,$$

where $C_*^{(1)} = (z_l - 1)/z_l$, *with*

$$z_l = \sup\left\{ z \;\middle|\; \sum_{k\in\mathbb{N}} [\rho_k]^{2l} \leq \frac{1}{z^2} \right\}.$$

4. Essence of the *j*-localization theorem

Before going into the detailed arguments for the j-localization theorem, we explain the point of Theorem 2, in this subsection. This will help us understanding what the new difficulty, beyond k-localization, is. Let $\gamma > H + 1$ and let ∂_λ be the differential operator with respect to λ. The main ingredient of the proof of the k-localization (in its simplest form)

$$\begin{aligned} r(n) &= C_H \int_{(0,\infty)} e^{in\lambda} |\hat{\psi}(\lambda)|^2 d\mu_H(\lambda) \\ &\leq \frac{C_H}{(1+n^2)^\gamma} \int_{(0,\infty)} |\hat{\psi}(\lambda)|^2 \, d\mu_H(\lambda) = \frac{\sigma_0^2}{(1+n^2)^\gamma}, \end{aligned} \tag{23}$$

for some $C_H > 0$ and $d\mu_H(\lambda) = \lambda^{-(1+2H)} d\lambda$, is a positive-definiteness argument, such as

$$0 \leq r(n) = \frac{C_H}{(1+n^2)^\gamma} \int_{(0,\infty)} e^{i\lambda n} \left[\iint_{W^2} \psi(s)\psi(t)(I - \partial_\lambda^2)^\gamma e^{i\lambda(s-t)} \lambda^{-(1+2H)} \, ds \, dt \right] d\lambda$$

and

$$\iint_{W^2} \psi(s)\psi(t)(I - \partial_\lambda^2)^\gamma e^{i\lambda(s-t)} \lambda^{-(1+2H)} \, ds \, dt \leq |\hat{\psi}(\lambda)|^2 \lambda^{-(1+2H)}$$

(Albeverio et al. [2]). Here the positive-definite functions are $(s,t) \mapsto e^{i\lambda(s-t)}$ and $(s,t) \mapsto \partial_\lambda^2[e^{i\lambda(s-t)}\lambda^{-(1+2H)}]$.

In the present case, in order to obtain the j-localization theorem, we will be reduced to evaluate

$$2^{-m/2} r(m,n) = C_H \int_{(0,\infty)} e^{i\lambda n} \hat{\psi}(\lambda) \overline{\hat{\psi}(2^m \lambda)}\, d\mu_H(\lambda) \tag{24}$$

with respect to $m \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$.

The difficulty in obtaining the desired evaluation is that the functions $(s, t) \mapsto e^{i\lambda(s-2^m t)}$ and $(s, t) \mapsto \partial_\lambda^2[e^{i\lambda(s-2^m t)}\lambda^{-(1+2H)}]$ are no longer nonnegative definite. In addition, we cannot take the absolute value inside the integral in (24): otherwise, the argument n on the right hand side disappears. Hence we cannot apply the same argument as in Albeverio et al. [2], at least directly. How can we evaluate (24)?

As a solution to this evaluation, we appeal to an argument that depends rather on a direct calculation in this paper. As a result, we indicate two factors $K_\gamma(m, n)$ and $\Psi_{\gamma,H}(m)$, as in Theorem 2. Here $K_\gamma(m, n)$ is to be compared with the right hand side of (23) and $\Psi_{\gamma,H}(m)$ is the factor that does not appear for the case of the k-localization ($m = j' - j = 0$) in (23).

5. Evaluations of $K_\gamma(m)$ and $\Psi_{\gamma,H}(m)$

In Subsections 5.1 and 5.2 below, we give evaluations of $K_\gamma(m, n)$ and $\Psi_{\gamma,H}(m)$ in Theorem 2 and $K_\gamma(m)$ in Theorem 3, respectively.

5.1. Evaluation of $K_\gamma(m,n)$ and $K_\gamma(m)$.

Let $q = (2\gamma - 1)l/2$.

Proposition 4. *Let $\gamma \in \mathbb{N}$ and $l \geq p$.*
(i) *For $m = 1$,*

$$K_\gamma^2(1, n) \leq \frac{\bar{c}_\gamma^{(K1)}}{(n+2)\{1+(n+2)^2\}^{2\gamma-1}}, \quad n \in \mathbb{N}_0 \tag{25}$$

where $\bar{c}_\gamma^{(K1)} = 3/5$ is sufficient. This leads to

$$K_\gamma^{(l)}(1) \leq c_\gamma^{(K1)} \left(\frac{3}{10}\right)^{l/2} \left(\frac{1}{5}\right)^q, \tag{26}$$

where $c_\gamma^{(K1)} = 1 + 10/(4q + l - 2)$ is sufficient.
(ii) *For $m = 2$,*

$$K_\gamma^2(2, n) \leq \frac{\bar{c}_\gamma^{(K2)}.}{\{1+(n+4)^2\}^{2\gamma-1}}, \quad n \in \mathbb{N}_0, \tag{27}$$

Table 1. Bounds for $K_\gamma^{(2)}(m)$, $m = 1, 2$.

(γ, l)	bounds for $K_\gamma^{(2)}(1)$	bounds for $K_\gamma^{(2)}(2)$
(1, 2)	1/5	1/30
(2, 1)	1/6	7/100
(1, 3)	1/20	3/200
(2, 2)	1/125	3 E−6

where $\bar{c}_\gamma^{(K2)} = 1/(32(2\gamma - 1))$ *is sufficient. This leads to*

$$K_\gamma^{(l)}(2) \le c_\gamma^{(K2)} \left(\frac{1}{17}\right)^q, \tag{28}$$

where $c_\gamma^{(K2)} = 1 + 17/(4q - 2)$ *is sufficient.*

For several values of parameters, the right hand sides of (26) and (28) are bounded numerically as in Table 1. Here $l = 2$ corresponds to the argument of the Hurst index estimation in the next section.

5.2. Evaluation of $\boldsymbol{\Psi_{\gamma,H}(m)}$. The second proposition is related to the evaluation of $\Psi_{\gamma,H}$.

Proposition 5. *For* $\gamma \in \mathbb{N}$

$$A_\gamma^2(1) \le \frac{1}{6}\left[1 + 3\left(\frac{1}{4}\right)^\gamma P_\gamma\left(\frac{3}{4}\right)\right] \triangleq \mathcal{A}_\gamma^2(1); \tag{29}$$

$$B_\gamma^2(1) \le \frac{1}{6}\left[1 + 2\left(\frac{1}{4}\right)^\gamma + 3\left(\frac{4}{9}\right)^\gamma \frac{P_\gamma(3/4)}{P_\gamma^2(1/4)}\right] \triangleq \mathcal{B}_\gamma^2(1), \tag{30}$$

where $P_\gamma(x) = \sum_{\nu=0}^{\gamma-1} \binom{\gamma-1+\nu}{\nu} x^\nu$ *for* $x \in [0, 1]$.

For $\gamma = 1$, the values of $\mathcal{A}_\gamma^2(1)$ and $\mathcal{B}_\gamma^2(1)$ are $\mathcal{A}_1^2(1) = 7/24 \fallingdotseq 0.2917$ and $\mathcal{B}_1^2(1) = 17/36 \fallingdotseq 0.4722$, by $P_1(x) \equiv 1$. These values of $\mathcal{A}_1^2(1)$ and $\mathcal{B}_1^2(1)$, not being sufficiently small, may cause a bad j-localization. So we may calculate more precise values of them directly as follows:

$$A_1^2(1) = \frac{1}{\pi}\int_0^\pi 4\sin^2\frac{\lambda}{4}\, d\lambda = \frac{16}{\pi}\int_0^{\pi/4}\left(\frac{1-\cos 2\lambda}{2}\right)^2 d\lambda = \frac{3}{2} - \frac{4}{\pi} \fallingdotseq 0.2268,$$

$$B_1^2(1) = \frac{4}{\pi}\int_0^{\pi/2} \frac{d\lambda}{(1+\cos\lambda)^2} = \frac{1}{\pi}\int_0^1 (1+y^2)\, dy = \frac{4}{3\pi} \fallingdotseq 0.4244,$$

Table 2. Bounds for $\Psi_{\gamma,H}(m)$, $m = 1, 2$.

γ	$\Psi_{\gamma,H}(1)$		$\Psi_{\gamma,H}(2)$	
	$H = 0.5$	$H = 1$	$H = 0.5$	$H = 1$
1	0.7059	1.779	0.3535	0.9490
2	0.5350	0.5754	5.562 E−2	7.800 E−2
3	0.4683	0.4693	2.964 E−2	3.082 E−2
4	0.4494	0.4494	2.387 E−2	2.392 E−2

where we have used in $B_1^2(1)$ a change of variable by $\tan(\lambda/2) = y$, so that $\cos\lambda = (1-y^2)/(1+y^2)$ and $d\lambda = [2/(1+y^2)]\,dy$.

As in this argument, obtaining sharper bounds of quantities in Theorems 2, 3 and Proposition 4 must be carried out carefully. Otherwise, the bounds as in tables here easily become so loose that the bounds are not useful. As in the proof below, we have contrived many ideas for calculations of the bounds.

5.3. Plugging-in the evaluations. An asymptotic evaluation of $r(m,n)$ is given as

$$2^{-m/2}r(m,n) = O(n^{-2(\gamma_0 - H)}), \quad n = |2^m k_1 - k_2|, \tag{31}$$

as $n \to \infty$, essentially (see e.g. [19]). The corresponding decay given by (20) is

$$2^{-m/2}r(m,n) \le c\Psi_{\gamma,H}(m) \times \begin{cases} [(n+2)\{1+(n+2)^2\}^{2\gamma-1}]^{-1}, & m = 1, \\ [1+(n+4)^2]^{-(2\gamma-1)}, & m = 2, \end{cases} \tag{32}$$

for $n \in \mathbb{N}$ and $c > 0$. We remark that the inequality in (32) holds for all $n \in \mathbb{N}$.

Apparently our estimation (32) is not better than (31). This is due to the fact that our estimation is based on the evaluation

$$r(n) \le \frac{\sigma_0^2}{(1+n^2)^{2\gamma}}, \quad \text{for all} \quad n \in \mathbb{N}_0, \tag{33}$$

obtained in Albeverio et al. [2], which itself implies $r(n) = O(n^{-2\gamma})$, a little bit worse *asymptotically* than (31). Here a trade-off is involved, however: (31) is precise but just an asymptotic evaluation and (33) is slightly worse, but still a useful pointwise evaluation.

The evaluation (32) of the cross-scale covariance $r(m,n)$ here involves summation in (21) and (22) with respect to $k \in \mathbb{N}$, so the resulting evaluation (32) is a little bit worse than (33) itself. However, as is shown in Albeverio et al. [2], the *pointwise* evaluation (33) has a great usefulness in the evaluation of the k-localization. We use the pointwise evaluation (32) for the cross-scale argument of j-localization as well, by the same reason: not just $r(m,n)$, but the evaluation such that the summation $R_H(m) = r^l(m,0) + 2\sum_{n\in\mathbb{N}} r^l(m,n)$ is sufficiently small, is necessary for our applications.

Table 3. Bounds for $R_H(m)/R_H$, $m = 1, 2$.

γ	$R_H(1)/R_H$		$R_H(2)/R_H$	
	$H = 0.5$	$H = 1$	$H = 0.5$	$H = 1$
1	2.527 E−1	3.145 E−1	1.826 E−2	3.289 E−2
2	9.304 E−4	5.379 E−4	4.157 E−7	2.012 E−7
3	2.108 E−5	1.058 E−5	2.924 E−10	7.904 E−11
4	6.653 E−7	3.327 E−7	5.612 E−13	1.409 E−13
5	2.284 E−8	1.142 E−8	1.356 E−15	3.391 E−16

Now we concatenate the above evaluations to obtain the estimate of $R_H(m)/R_H = [\Psi_{\gamma,H}(m)K_\gamma(m)]^l/C_*^{(1)}$. From Theorem 3, Propositions 4 and 5, we have

$$\frac{R_H(1)}{R_H} \le C^{(R1)}\left[\Psi_{\gamma,H}(1)\left\{\frac{3}{10}\left(\frac{1}{5}\right)^{2\gamma-1}\right\}^{1/2}\right]^l \tag{34}$$

with $C^{(R1)} = c_\gamma^{(K1)}/C_*^{(1)}$ and

$$\frac{R_H(2)}{R_H} \le C^{(R2)}\left[\Psi_{\gamma,H}(2)\left(\frac{1}{17}\right)^{(2\gamma-1)/2}\right]^l \tag{35}$$

with $C^{(R2)} = c_\gamma^{(K2)}/C_*^{(1)}$.

Numerical evaluations for the bound on the right hand sides of (34) and (35), for $\gamma = 1$ to 5 are given in Table 3. Here we have set $l = 2$, which is the case of Hurst index estimation in the next section. The ratios $R_H(m)/R_H$ become quite small for $\gamma \ge 2$, as can be seen.

6. Main result (2)—application of the j-localization to the Hurst index estimation

In this section, we apply the j-localization property to the problem of the wavelet-based Hurst index estimation for FBM. Especially, it turns out that the evaluation of $R_H(m)/R_H$ works effectively in determining the scale upper bound that achieves the minimum variance of the estimator.

The wavelet-based method was proposed by Abry et al. [1]. The method is based on the variance $\mathrm{Var}[s_j(1)] = \sigma_{H,j}^2 = \theta_j$, $j = 1, \dots, J$ of the wavelet coefficient of FBM at scale $j \in \mathcal{J} = \{1, \dots, J\}$. We may write here $\xi_{\mathcal{J}} \equiv \xi_J$ and $v_{H,\mathcal{J}}^2 = v_{H,J}^2$. Then $\sigma_{H,j}^2$ is estimated consistently by $\hat\theta_{j,T}$ with $g(x) = x^2$ and the estimator, denoted by $\hat H_T$, is given by

$$\hat H_{J,T} = \sum_{j=1}^{J} a_j\left[\log_2\hat\theta_{j,T} - \frac{1}{J}\sum_{j=1}^{J}\log_2\hat\theta_{j,T}\right] - \frac{1}{2} = f(\hat{\boldsymbol\theta}_{J,T}),$$

where $\{a_j;\ j=1,\dots,J\}$ is the linear least square regression coefficient given by $a_j = (x_j - \bar{x}_J)/\big[2\sum_{j=1}^{J}(x_j - \bar{x}_J)^2\big]$ with $x_j = j$ and $\bar{x}_J = J^{-1}\sum_{j=1}^{J} x_j = (J+1)/2$. Here $\varphi_j(x)$ in (9) is $\log_2 x$. By an elementary calculation, $a_j = [6j - 3(J+1))]/[(J-1)J(J+1)]$.

If $\gamma_0 \ge 2$ (actually if $\gamma_0 > H + (1/2p) = H + (1/4)$), then we can rewrite the modified CLT in (10) and (11) as follows. Here, $\dot{\varphi}_j(\theta_j) = 1/\theta_j = 1/\sigma_{H,j}^2$. Also,

$$\hat{\theta}_{j,T} - \theta_j = \frac{\theta_j}{N_{J,T}} \sum_{n=1}^{N_{J,T}} \left[\frac{1}{d_j} \sum_{k=1}^{d_j} H_2(s_j(d_j n + k)) \right]$$

and hence, in the present case, $g(z_j) - \theta_j = \sigma_{H,j}^2 H_2(z_j/\sigma_{H,j}) = \sigma_{H,j}^2[(z_j/\sigma_{H,j})^2 - 1]$ and $l = p = 2$. Moreover, we have $(\hat{\theta}_{j,T} - \theta_j)\cdot\dot{\varphi}_j(\theta_j) = N_{J,T}^{-1}\sum_{n=1}^{N_{J,T}}\big[(1/d_j)\sum_{k=1}^{d_j} H_2(s_j(d_j n + k))\big]$. In this case, $\xi_J(m)$ in (13) reduces just to $\xi_J(m) = \sum_{j=1}^{J-m} 2^j a_j a_{j+m}$.

The following Proposition gives the evaluation of the limiting variance $v_{H,J}^2$ in terms of diagonal component $\xi_J R_H^2$ in (15), for the case of the Hurst index estimation.

Proposition 6. *Let $\gamma > H + (1/2)$ and let the wavelet ψ be associated with an MRA. Then the limiting variance $v_{H,J}^2$ satisfies the following evaluation*:

$$C_{H,J}\xi_J R_H^2 \le v_{H,J}^2 \le \xi_J R_H^2 \quad \textit{for} \quad J = 2, 3 \tag{36}$$

with $C_{H,2} = 1 - (2^{1-2H}/3)$ and $C_{H,3} = 1 - (2^{1-4H}/5)$;

$$\xi_J R_H^2 \le v_{H,J}^2 \le C_{H,J}\xi_J R_H^2 \quad \textit{for} \quad J \ge 4; \tag{37}$$

For $C_{H,J}$, $J \ge 4$, it is enough to take

$$\begin{aligned} C_{H,J} &= 1 + 2\cdot\frac{R_H^2(1)}{R_H^2}\cdot Z_{H,J}, \quad \textit{with} \\ Z_{H,J} &= \frac{2^{-(2H+1)}}{1 - 2^{-(2H+1)}}\cdot\frac{2^J(J^2 - 8J + 23) - 2(J^2 + 4J + 11)}{2^J(J^2 - 6J + 17) - (J^2 + 6J + 17)}. \end{aligned} \tag{38}$$

Hence, for $C_{H,J}$, $J \ge 4$, we can write

$$C_{H,J} = 1 + \frac{2^{-2H}}{1 - 2^{-(2H+1)}}\cdot\frac{1}{C_*^{(1)}}[\Psi_{\gamma,H}(1)K_{\gamma,H}(1)]^2\cdot(1 + O(J^{-1}))$$

as $J \to \infty$, where $K_{\gamma,H}(1) \equiv K_{\gamma,H}^{(l)}(1)|_{l=2}$.

REMARK 1. For the ordering of $v^2_{H,J}$, $J \ge 4$, we should remark that, for $\xi_J(m)$ given by $\xi_J(m) = \sum_{j=1}^{J-m} 2^j a_j a_{j+m}$, it is positive for $1 \le m \le \lfloor (J-2)/2 \rfloor$ and negative for $\lfloor (J-2)/2 \rfloor + 1 \le m \le J-1$, which follows by induction.

We can apply Proposition 6 to evaluate the ordering of $v^2_{H,J}$, $J = 2, 3, \ldots$ precisely.

Theorem 7. *For all* $0 < H < 1$,

$$v^2_{H,2} > v^2_{H,3} > v^2_{H,4} > v^2_{H,5} < v^2_{H,6} < v^2_{H,7} < \cdots ;$$

Thus, $\min_{J \ge 2} v^2_{H,J} = v^2_{H,5}$.

7. Proof of Theorem 2

We will prove Theorem 2 for the case of $m = 1$ and $m = 2$ separately in Subsections 7.1 and 7.2 below, respectively.

For the evaluation of $r(m, n)$, we have only to prove the case of $n \ge 0$, since $r(m, n) = r(m, -n)$ by Lemma 8 below. As in Albeverio et al. [2, Theorem 1], we can write $r(m, n) = \mathrm{Cov}[s_0(k_0 + k_1), s_m(k_0 + k_2)]|_{2^m k_2 - k_1 = n}$ as (recall the initial shift k_0 in (4))

$$\begin{aligned} r(m, n) &= C_H 2^{m/2} \int_{(0,\infty)} \hat{\psi}_{0,k_0+k_1}(\lambda) \overline{\hat{\psi}_{m,k_0+k_2}(\lambda)}\, d\mu_H(\lambda) \\ &= C_H 2^{m/2} \int_{(0,\infty)} \exp[i\lambda(n + (2^m - 1)k_0)] \hat{\psi}(\lambda) \overline{\hat{\psi}(2^m\lambda)}\, d\mu_H(\lambda), \end{aligned}$$

with $d\mu_H(\lambda) = \lambda^{-(2H+1)}\, d\lambda$. This can be rewritten as

$$2^{-m/2} r(m, n) = C_H \int_{\mathbb{R}\setminus 0} \exp[i\lambda(n + (2^m - 1)k_0)] \hat{\psi}(\lambda) \overline{\hat{\psi}(2^m\lambda)}\, d\mu_H(\lambda), \tag{39}$$

by the argument in the proof of Lemma 8, where μ_H is redefined by $d\mu_H(\lambda) = |\lambda|^{-(2H+1)} d\lambda$ on $\mathbb{R} \setminus 0$. We will evaluate the Fourier integral in (39) below.

7.1. The case of $m = 1$. First, let us consider the case of $m = 1$. Set $G_1 = \bigcup_{l \in \mathbb{Z}} \{[\pi, 3\pi] + 4\pi l\}$ and $G_2 = \bigcup_{l \in \mathbb{Z}} \{[-\pi, \pi] \setminus 0 + 4\pi l\}$, respectively.

We divide the Fourier integral into two parts as

$$\left(\int_{G_1} + \int_{G_2} \right) e^{i\lambda(n+k_0)} \hat{\psi}(\lambda) \overline{\hat{\psi}(2\lambda)}\, d\mu_H(\lambda) \triangleq I_1(m, n) + I_2(m, n)|_{m=1}.$$

We recall the expression for the MRA wavelet ψ in (19). Then, the 4π-periodic functions $\overline{\hat{\psi}(2\lambda)} \big/ \overline{\hat{\psi}(\lambda)}$ and $\hat{\psi}(\lambda)/\hat{\psi}(2\lambda)$ turn out to be bounded, with absolute values less than or equal to 1, on G_1 G_2, respectively (see Remark 6 below).

Let us consider $I_1(1,n)$ first. Subdivide it into the integral on $G_{11} = \bigcup_{l\in\mathbb{Z}}[\pi, 2\pi] + 4\pi l$ and $G_{12} = \bigcup_{l\in\mathbb{Z}}[2\pi, 3\pi] + 4\pi l$. Due to the fact that $|\hat{\psi}(2\lambda)/\hat{\psi}(\lambda)|$ is symmetric about $\lambda = 2\pi$, we can apply the same upper bound for the two subdivided integrals, as seen below.

On $[\pi, 2\pi]$, it turns out that the first term on the right hand side of

$$\hat{\psi}(\lambda)\overline{\hat{\psi}(2\lambda)} = \frac{\overline{\hat{\psi}(2\lambda)}}{\overline{\hat{\psi}(\lambda)}} \cdot |\hat{\psi}(\lambda)|^2$$

is bounded with absolute values less than or equal to 1. We can write

$$\frac{\overline{\hat{\psi}(2\lambda)}}{\overline{\hat{\psi}(\lambda)}} = e^{i\lambda/2}\frac{m_0(\lambda+\pi)\overline{m_0(\lambda/2)}}{m_0(\lambda/2+\pi)} = e^{-i\lambda(N_2-N_1-1)/2}\varphi_\gamma^{(1)}, \tag{40}$$

for some function $\varphi_\gamma^{(1)}$ with the Fourier series expansion $\varphi_\gamma^{(1)}(\lambda) = \sum_{k\in\mathbb{N}} \alpha_k^{(1)} e^{i2k\lambda}$, $\lambda \in [\pi, 2\pi]$ by Lemma 9 below. Hence we have

$$\begin{aligned}\int_{G_{11}} e^{i\lambda(n+k_0)}\hat{\psi}(\lambda)\overline{\hat{\psi}(2\lambda)}\, d\mu_H(\lambda) &= \int_{G_{11}} e^{i\lambda(n+\zeta_1)}\varphi_\gamma^{(1)}(\lambda)|\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda)\\ &= \sum_{k\in\mathbb{N}} \alpha_k^{(1)} \int_{G_{11}} e^{i\lambda(n+\zeta_1+2k)}|\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda),\end{aligned}$$

where

$$\zeta_1 \triangleq k_0 - \frac{N_2-N_1-1}{2} = \begin{cases} 0 & \text{for} \quad N_2-N_1\text{: odd,}\\ \frac{1}{2} & \text{for} \quad N_2-N_1\text{: even}\end{cases} \tag{41}$$

(see (5)). From the fact that

$$r(n) = C_H \int_{\mathbb{R}} e^{i\lambda n}|\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda) \le \frac{\sigma_{H,0}^2}{(1+n^2)^\gamma}, \quad n \in \mathbb{N}$$

by the proof of [2, Theorem 1] (recall that $\sigma_{H,0}^2 = C_H \int_{\mathbb{R}} |\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda)$), it then follows

$$\begin{aligned}&\sum_{k\in\mathbb{N}} \alpha_k^{(1)} \int_{G_{11}} e^{i\lambda(n+\zeta_1+2k)}|\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda)\\ &\le \sum_{k\in\mathbb{N}} |\alpha_k^{(1)}| \cdot \frac{\int_{G_{11}}|\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda)}{[1+(n+\zeta_1+2k)^2]^\gamma}\\ &\le [A_\gamma(1)K_\gamma(1, n+\zeta_1)] \int_{G_{11}} |\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda),\end{aligned} \tag{42}$$

where $A_\gamma(m)|_{m=1}$ and $K_\gamma(m,\ \cdot)|_{m=1}$ are such that

$$A_\gamma^2(1) = \frac{1}{\pi}\sum_{k\in\mathbb{N}}|\alpha_k^{(1)}|^2 = \frac{1}{\pi}\int_\pi^{2\pi}\left|\frac{\hat\psi(2\lambda)}{\hat\psi(\lambda)}\right|^2 d\lambda$$

and

$$K_\gamma^2(1,n) = \sum_{k\in\mathbb{N}}\{1+(n+2k)^2\}^{-2\gamma},$$

respectively. By the reason stated above, we also have, for $[2\pi, 3\pi]$, a similar inequality for the integral with $[\pi, 2\pi]$ replaced by $[2\pi, 3\pi]$. Hence,

$$\begin{aligned} I_1(1,n) &\leq [A_\gamma(1)K_\gamma(1, n+\zeta_1)]\int_{G_1}|\hat\psi(\lambda)|^2 d\mu_H(\lambda) \\ &\leq [A_\gamma(1)K_\gamma(1, n+\zeta_1)]\sigma_{H,0}^2. \end{aligned} \tag{43}$$

Similarly, we consider $I_2(1,n)$ on subdivided intervals $G_{21} = \bigcup_{l\in\mathbb{Z}}(0,\pi]+4\pi l$ and $G_{22} = \bigcup_{l\in\mathbb{Z}}[-\pi, 0)+4\pi l$, and the two evaluations turn out to have upper bounds that have the same form except for the intervals of the Fourier integrals. So, for $\bigcup_{l\in\mathbb{Z}}(0,\pi]+4\pi l$, we can write

$$\frac{\hat\psi(\lambda)}{\hat\psi(2\lambda)} = e^{i\lambda(N_2-N_1+1)/2}\varphi_\gamma^{(2)}, \quad \text{with} \quad \varphi_\gamma^{(2)} = \sum_{k\in\mathbb{N}}\alpha_k^{(2)}e^{i2k\lambda}.$$

Thus we have

$$\begin{aligned} &\int_{G_{21}} e^{i\lambda(n+k_0)}\hat\psi(\lambda)\overline{\hat\psi(2\lambda)}\,d\mu_H(\lambda) \\ &= \int_{G_{21}} e^{i(n+\bar\zeta_1)\lambda}\varphi_\gamma^{(2)}(\lambda)|\hat\psi(2\lambda)|^2\,d\mu_H(\lambda) \\ &\leq B_\gamma(1)\cdot K_\gamma(1, n+\bar\zeta_1)\cdot\int_{G_{21}}|\hat\psi(2\lambda)|^2\,d\mu_H(\lambda) \\ &= [B_\gamma(1)K_\gamma(1,n+\bar\zeta_1)]2^{2H}\int_{\bigcup_{l\in\mathbb{Z}}(0,2\pi]+8\pi l}|\hat\psi(\lambda)|^2\,d\mu_H(\lambda), \end{aligned}$$

where

$$\bar\zeta_1 \triangleq k_0 + \frac{N_2-N_1+1}{2} = \begin{cases} N_2-N_1\ (=2\gamma-1) & \text{for} \quad N_2-N_1\text{: odd}, \\ N_2-N_1+\dfrac{1}{2}\ (=2\gamma-1/2) & \text{for} \quad N_2-N_1\text{: even}, \end{cases}$$

and $B_\gamma(1)$ is such that

$$B_\gamma^2(1) = \frac{1}{\pi}\sum_{k\in\mathbb{N}}|\alpha_k^{(2)}|^2 = \frac{1}{\pi}\int_{(0,\pi]}\left|\frac{\hat\psi(\lambda)}{\hat\psi(2\lambda)}\right|^2 d\lambda.$$

A similar inequality holds for the integral with G_{21} replaced by G_{22} as well. Hence, writing $2G_2 = \bigcup_{l\in\mathbb{Z}}[-2\pi, 2\pi] \setminus 0 + 8\pi l$, we have

$$
\begin{aligned}
I_2(1, n) &\le 2^{2H}[B_\gamma(1)K_\gamma(1, n + \bar{\zeta}_1)] \int_{2G_2} |\hat{\psi}(\lambda)|^2 \, d\mu_H(\lambda) \\
&\le 2^{2H}\left[B_\gamma(1) \left\{ \frac{5}{(1+2\gamma)^2} \right\}^\gamma K_\gamma(1, n + \zeta_1) \right] \sigma^2_{H,0},
\end{aligned} \tag{44}
$$

where we have used the comparison of $K^2_\gamma(1, n + \bar{\zeta}_1)$ with $K^2_\gamma(1, n + \zeta_1)$:

$$
\begin{aligned}
K^2_\gamma(1, n + \bar{\zeta}_1) &\le \frac{3/[2(2\gamma+1)]}{(n + 2\gamma + 1)\{1 + (2\gamma+1)^2\}^{2\gamma-1}} \\
&\le \left\{ \frac{5}{(1+2\gamma)^2} \right\}^{2\gamma} \cdot K^2_\gamma(1, n + \zeta_1).
\end{aligned}
$$

The first inequality of this is obtained by the argument in the proof of Proposition 4. Combining (43) and (44) and taking $\zeta_1 = 0$ results in (20) for $m = 1$. $\square$

REMARK 2. The initial shift k_0 in (4) is essential here in making the upperbound in (42) a reasonable one. In fact, the leading term in the resulting upperbound for $K_\gamma(1, n)$ is given by $\zeta_1 = 0$ (see (41) and (42)). The indices n and ζ_1 start from 0, and k from 1. If one or more of the three would start from negative indices, then the resulting upperbound of $K_\gamma(m)$ would become considerably worse.

We have divided the Fourier integral (39) into those on the intervals $(0, \pi]$ and $(\pi, 2\pi]$, and not simply $[-\pi, \pi] \setminus 0$ and $(\pi, 3\pi]$. The reason for this will be clear in the proof of the case of $m = 2$ below.

Here a question may arise. The smaller we divide the intervals, the sharper will the resulting evaluation be? The answer is not clear presently. The two factors $K_\gamma(1,n)$ and $\Psi_{\gamma,H}(1)$ themselves will be smaller indeed, but we have to sum them, as we did in

$$
\begin{aligned}
&I_1(1, n) + I_2(1, n) \\
&\le K_\gamma(1, n + \zeta_1)\left\{ A_\gamma(1) \int_{G_1} |\hat{\psi}(\lambda)|^2 \, d\mu_H(\lambda) \right. \\
&\qquad\qquad \left. + 2^{2H} B_\gamma(1) \left[\frac{5}{(1+2\gamma)^2} \right]^\gamma \int_{2G_2} |\hat{\psi}(\lambda)|^2 \, d\mu_H(\lambda) \right\};
\end{aligned}
$$

If the intervals of the two integrals were not overlapping, then the answer would be YES with the upper bound for $I_1(1, n) + I_2(1, n)$ simply given by

$$
K_\gamma(1, n + \zeta_1) \int_{\mathbb{R}\setminus 0} |\hat{\psi}(\lambda)|^2 \, d\mu_H(\lambda) \cdot \max\left\{ A_\gamma(1), 2^{2H} B_\gamma(1) \left[\frac{5}{(1+2\gamma)^2} \right]^\gamma \right\}.
$$

However, this is not true. Thus, if we divide the original integral in (39) into those on smaller intervals $I_1(1, n) + \cdots + I_N(1, n)$ with $N > 2$, then we have to take the summation of the corresponding N upper bounds. To find an optimal way of doing the division might be of interest. This argument holds for $m \geq 2$ as well. From a realistic point of view, theoretical statements may be taken through a trade-off between precision of evaluations and statement's simplicity.

7.2. The case of $m = 2$. In (39), let us take $m = 2$. We can consider only $\lambda \in (0, 2\pi]$ as in the case of $m = 1$, because of the symmetry of the functions $|\hat{\psi}(\lambda)/\hat{\psi}(2\lambda)|\mathbb{I}_{|\lambda|\leq\pi}$ and $|\hat{\psi}(2\lambda)/\hat{\psi}(\lambda)|\mathbb{I}_{\pi\leq|\lambda|\leq 3\pi}$ or their scaled ones, about $\lambda = 0$.

Divide the Fourier integral into four parts as

$$C_H \int_{\mathbb{R}\setminus 0} e^{i\lambda(n+3k_0)}\hat{\psi}(\lambda)\overline{\hat{\psi}(4\lambda)}\, d\mu_H(\lambda) = \sum_{\nu=1}^{4} I_\nu(m, n)\bigg|_{m=2},$$

where

$$I_\nu(2, n) = \int_{G_\nu^{(2)}} e^{i\lambda(n+3k_0)}\hat{\psi}(\lambda)\overline{\hat{\psi}(4\lambda)}\, d\mu_H(\lambda),$$

and where $G_\nu^{(2)} = \bigcup_{l\in\mathbb{Z}}((\nu-1)\pi/2, \nu\pi/2] + 4\pi l$.

For $\nu = 1$, we have that the first two terms on the right hand side of

$$\hat{\psi}(\lambda)\overline{\hat{\psi}(4\lambda)} = \frac{\hat{\psi}(\lambda)}{\hat{\psi}(2\lambda)} \cdot \frac{\hat{\psi}(2\lambda)}{\hat{\psi}(4\lambda)} \cdot |\hat{\psi}(4\lambda)|^2$$

are bounded with absolute values less than or equal to 1. We can write

$$\hat{\psi}(\lambda)/\hat{\psi}(2\lambda) = e^{i(N_2-N_1+1)/2}\varphi_1^{(m,\nu)}(\lambda)|_{m=2,\nu=1} \quad \text{on} \quad (0, \pi/2]$$

for a function $\varphi_1^{(2,1)}(\lambda)$, which is expanded in a Fourier series on $(0,\pi/2]$ as $\varphi_1^{(2,1)}(\lambda) = \sum_{k\in\mathbb{N}} \beta_{1,k}^{(2,1)} e^{i4k\lambda}$. Similarly we can write

$$\hat{\psi}(2\lambda)/\hat{\psi}(4\lambda) = e^{i(N_2-N_1+1)}\varphi_2^{(2,1)}(2\lambda) \quad \text{on} \quad (0, \pi/2]$$

and expand the function $\varphi_2^{(2,1)}(\lambda) = \varphi_1^{(2,1)}(2\lambda)$ on $(0,\pi/2]$ as $\varphi_2^{(2,1)}(\lambda) = \sum_{k\in\mathbb{N}} \beta_{2,k}^{(2,1)} e^{i4k\lambda}$. These $\beta_{1,k}^{(2,1)}$ and $\beta_{2,k}^{(2,1)}$ are such that

$$\begin{aligned}
[B_{1,\gamma}^{(2,1)}]^2 &\triangleq \frac{1}{\pi/2}\sum_{k\in\mathbb{N}}|\beta_{1,k}^{(2,1)}|^2 = \frac{1}{\pi/2}\int_{(0,\pi/2]}\left|\frac{\hat{\psi}(\lambda)}{\hat{\psi}(2\lambda)}\right|^2 d\lambda,\\
[B_{2,\gamma}^{(2,1)}]^2 &\triangleq \frac{1}{\pi/2}\sum_{k\in\mathbb{N}}|\beta_{2,k}^{(2,1)}|^2 = \frac{1}{\pi/2}\int_{(0,\pi/2]}\left|\frac{\hat{\psi}(2\lambda)}{\hat{\psi}(4\lambda)}\right|^2 d\lambda = \frac{1}{\pi}\int_{(0,\pi]}\left|\frac{\hat{\psi}(\lambda)}{\hat{\psi}(2\lambda)}\right|^2 d\lambda\\
&= [B_\gamma^{(1)}]^2,
\end{aligned}$$

respectively. Then, writing $4G_1^{(2)} = \bigcup_{l\in\mathbb{Z}}(0, 2\pi] + 16\pi l$, we have

$$\begin{aligned} I_1(2, n) &= \int_{G_1^{(2)}} e^{i\lambda(n+3k_0)}|\hat{\psi}(4\lambda)|^2 \varphi_1^{(2,1)}(\lambda)\varphi_2^{(2,1)}(\lambda)\, d\mu_H(\lambda) \\ &= \sum_{k_1\in\mathbb{N}} \beta_{1,k_1}^{(2,1)} \sum_{k_2\in\mathbb{N}} \beta_{2,k_2}^{(2,1)} \int_{G_1^{(2)}} e^{i\lambda(n+\bar{\zeta}_{2,1}+4k_1+4k_2)}|\hat{\psi}(4\lambda)|^2\, d\mu_H(\lambda) \\ &\le [B_\gamma(1)B_\gamma(2, 1)K_\gamma(2, n+\bar{\zeta}_{2,1})]4^{2H} \int_{4G_1^{(2)}} |\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda), \end{aligned}$$

where $K_\gamma(2, n)$ and $\bar{\zeta}_{2,1}$ are such that

$$[K_\gamma(2, n)]^2 = \sum_{k_1\in\mathbb{N}} \sum_{k_2\in\mathbb{N}} \{1 + (n + 4k_1 + 4k_2)^2\}^{-2\gamma}$$

and

$$\bar{\zeta}_{2,1} = 3k_0 + \frac{3(N_2 - N_1 + 1)}{2} = \begin{cases} 3(N_2 - N_1)\ (= 6\gamma - 3) & N_2 - N_1\text{: odd,} \\ 3(N_2 - N_1) + \dfrac{3}{2}\ (= 6\gamma - 3/2) & N_2 - N_1\text{: even.} \end{cases}$$

Similarly, for $\nu = 2$, writing $2G_2^{(2)} = \bigcup_{l\in\mathbb{Z}}(\pi, 2\pi] + 8\pi l$, we have

$$\begin{aligned} I_2(2, n) &= \int_{G_2^{(2)}} e^{i\lambda(n+3k_0)}|\hat{\psi}(2\lambda)|^2 \overline{\frac{\hat{\psi}(4\lambda)}{\hat{\psi}(2\lambda)}}\, \frac{\hat{\psi}(\lambda)}{\hat{\psi}(2\lambda)}\, d\mu_H(\lambda) \\ &\le [A_\gamma(1)B_\gamma(2, 2)K_\gamma(2, n+\bar{\zeta}_{2,2})]2^{2H} \int_{2G_2^{(2)}} |\hat{\psi}(\lambda)|^2\, d\mu_H(\lambda), \end{aligned}$$

for functions $\hat{\psi}(4\lambda)/\hat{\psi}(2\lambda)$ and $\hat{\psi}(\lambda)/\hat{\psi}(2\lambda)$ on $\bigcup_{l\in\mathbb{Z}}(\pi/2, \pi] + 4\pi l$ bounded by or equal to 1, where

$$\begin{aligned} \bar{\zeta}_{2,2} &= 3k_0 - \frac{2(N_2 - N_1 - 1)}{2} + \frac{N_2 - N_1 + 1}{2} \\ &= \begin{cases} N_2 - N_1\ (= 2\gamma - 1) & N_2 - N_1\text{: odd,} \\ N_2 - N_1 + \dfrac{3}{2}\ (= 2\gamma + 1/2) & N_2 - N_1\text{: even} \end{cases} \end{aligned}$$

and

$$[B_\gamma(2, 2)]^2 = \frac{1}{\pi/2} \sum_{k\in\mathbb{N}} |\beta_k^{(2,2)}|^2 = \frac{1}{\pi/2} \int_{(\pi/2,\pi]} \left|\frac{\hat{\psi}(\lambda)}{\hat{\psi}(2\lambda)}\right|^2 d\lambda.$$

Also, for $\nu = 3$,

$$
\begin{aligned}
I_3(2, n) &= \int_{G_3^{(2)}} e^{i\lambda(n+3k_0)} |\hat{\psi}(\lambda)|^2 \frac{\overline{\hat{\psi}(4\lambda)}}{\overline{\hat{\psi}(2\lambda)}} \frac{\overline{\hat{\psi}(2\lambda)}}{\overline{\hat{\psi}(\lambda)}} \, d\mu_H(\lambda) \\
&\leq [A_\gamma(1) A_\gamma(2, 3) K_\gamma(2, n + \bar{\zeta}_{2,3})] \int_{G_3^{(2)}} |\hat{\psi}(\lambda)|^2 \, d\mu_H(\lambda),
\end{aligned}
$$

where $\bar{\zeta}_{2,3} \equiv \zeta_2$ (this $\bar{\zeta}_{2,3}$ will be the "basis" among $\bar{\zeta}_{2,\nu}$, $\nu = 1, 2, 3, 4$ corresponding to the leading term of $K_\gamma(2, n + \bar{\zeta}_{2,\nu})$) is given by

$$
\zeta_2 = 3k_0 - \frac{3(N_2 - N_1 - 1)}{2} = \begin{cases} 0 & N_2 - N_1\colon \text{ odd}, \\ \dfrac{3}{2} & N_2 - N_1\colon \text{ even}, \end{cases}
$$

and where

$$
[A_\gamma(2, 3)]^2 \triangleq \frac{1}{\pi/2} \sum_{k \in \mathbb{N}} |\alpha_k^{(2)}|^2 = \frac{1}{\pi/2} \int_{(\pi, 3\pi/2]} \left| \frac{\hat{\psi}(2\lambda)}{\hat{\psi}(\lambda)} \right|^2 d\lambda.
$$

Finally, for $\nu = 4$,

$$
\begin{aligned}
I_4(2, n) &= \int_{G_4^{(2)}} e^{i\lambda(n+3k_0)} |\hat{\psi}(\lambda)|^2 \frac{\overline{\hat{\psi}(4\lambda)}}{\overline{\hat{\psi}(\lambda)}} \, d\mu_H(\lambda) \\
&= \sum_{k \in \mathbb{N}} \alpha_k^{(2,4)} \int_{G_4^{(2)}} e^{i\lambda(n+\zeta_{2,4}+4k)} |\hat{\psi}(\lambda)|^2 \, d\mu_H(\lambda) \\
&\leq [A_\gamma(2, 4) \bar{K}_\gamma(2, n + \bar{\zeta}_{2,4})] \int_{G_4^{(2)}} |\hat{\psi}(\lambda)|^2 \, d\mu_H(\lambda),
\end{aligned}
$$

where $A_\gamma(2, 4)$, $\bar{K}_\gamma(2, n)$ and $\bar{\zeta}_{2,4}$ are given by

$$
[A_\gamma(2, 4)]^2 = \frac{1}{\pi/2} \sum_{k \in \mathbb{N}} |\alpha_k^{(2,4)}|^2 = \frac{1}{\pi/2} \int_{3\pi/2}^{2\pi} \left| \frac{\hat{\psi}(4\lambda)}{\hat{\psi}(\lambda)} \right|^2 d\lambda,
$$

$\bar{K}_\gamma(2, n) = \sum_{k \in \mathbb{N}} \{1 + (n + 4k)^2\}^{-2\gamma}$ and

$$
\bar{\zeta}_{2,4} = 3k_0 + \frac{3}{2} = \begin{cases} \dfrac{3(N_2 - N_1)}{2} \; (= 3\gamma - 3/2) & N_2 - N_1\colon \text{ odd}, \\ \dfrac{3(N_2 - N_1 + 1)}{2} \; (= 3\gamma) & N_2 - N_1\colon \text{ even}, \end{cases}
$$

respectively. Comparing the evaluations of $K_\gamma(2, n+\bar{\zeta}_{2,\nu})$ for $\nu = 1, 2$ and $\bar{K}_\gamma(2, n+\bar{\zeta}_{2,4})$ with $K_\gamma(2, n+\zeta_2)$ based on (27), we can combine the evaluations of $I_\nu(2, n)$ for $\nu = 1$ to 4, to obtain the upper bound of $r(2, n)$,

$$
\begin{aligned}
K_\gamma(2, n+\zeta_2)\Bigg[& B_\gamma(1) B_\gamma(2, 1)\left\{\frac{17}{1+(6\gamma+1)^2}\right\}^{(2\gamma-1)/2} \frac{4^{2H}}{\sqrt{32(2\gamma-1)}} \int_{4G_1^{(2)}} |\hat{\psi}(\lambda)|^2 \, d\mu_H \\
& + A_\gamma(1) B_\gamma(2, 2)\left\{\frac{17}{1+(2\gamma+3)^2}\right\}^{(2\gamma-1)/2} \frac{2^{2H}}{\sqrt{32(2\gamma-1)}} \int_{2G_2^{(2)}} |\hat{\psi}(\lambda)|^2 \, d\mu_H \\
& + A_\gamma(1) A_\gamma(2, 3) \frac{1}{\sqrt{32(2\gamma-1)}} \int_{G_3^{(2)}} |\hat{\psi}(\lambda)|^2 d\mu_H \\
& + A_\gamma(2, 4)\sqrt{\frac{6}{30\gamma+25}}\left\{\frac{17}{1+(3\gamma+3/2)^2}\right\}^{(2\gamma-1)/2} \int_{G_4^{(2)}} |\hat{\psi}(\lambda)|^2 \, d\mu_H \Bigg].
\end{aligned}
$$

It turns out by computation that the coefficient of the integral of the third term in the square brackets is greater than that of the fourth term for all $\gamma \in \mathbb{N}$. Hence, gathering the third and fourth terms, and rewriting $A_\gamma(2, 3) \equiv A_\gamma(2)$, results in (20). □

Lemma 8. *For each $m = 1, 2, \ldots, r(m, -\nu) = r(m, \nu)$, $\nu \in \mathbb{Z}$.*

Lemma 9. *The Fourier coefficients $\{\alpha_k^{(1,2)}\}$ and $\{\beta_k^{(1,1)}\}$ vanish for $k \in -\mathbb{N}_0$.*

REMARK 3. For $\nu = 4$, the reason why we do not follow (40) but take

$$
\hat{\psi}(\lambda)\overline{\hat{\psi}(4\lambda)} = |\hat{\psi}(\lambda)|^2 \frac{\overline{\hat{\psi}(4\lambda)}}{\overline{\hat{\psi}(\lambda)}}
$$

is because the value of the integral $\int_{(3\pi/2, 2\pi]} |\hat{\psi}(2\lambda)/\hat{\psi}(4\lambda)|^2 \, d\lambda$ can be large, since $|\hat{\psi}(2\lambda)/\hat{\psi}(4\lambda)| \geq 1$ on $(3\pi/2, 2\pi]$.

8. Proof of Theorems 3 and 7

8.1. Proof of Theorem 3. Let $r(k) = \mathrm{Cov}[X_k, X_0]$ for a general stationary sequence $\{X_k;\, k = 1, \ldots, N\}$ and let $\Sigma^{(1)}$ be its covariance matrix. Then, from Albeverio et al. [2, Proposition 1 and Theorem 2], we have the following statement:

$$
\text{(45)} \quad \text{if} \quad \sum_{k \in \mathbb{N}} \left\{\frac{r^l(k)}{r^l}\right\}^2 \leq \frac{1}{\zeta_l^2} \quad \text{for some} \quad \zeta_l > 1 \quad \text{then} \quad C_{l*}\Lambda^l \leq \Sigma^{(1)} \leq C_l^* \Lambda^l,
$$

where $C_{l*} = (\zeta_l - 1)/\zeta_l$, $C_l^* = (\zeta_l + 1)/\zeta_l$, and $\Lambda = \mathrm{diag}(r, \ldots, r)$ with $r \equiv r(0) = \mathrm{Var}[X_1]$. This statement implies

$$C_*^{(1)} r^l \leq \sum_{n\in\mathbb{Z}} r^l(n) = r^l + 2\sum_{n\in\mathbb{N}} r^l(n). \tag{46}$$

In fact,

$$r^l + 2\sum_{n\in\mathbb{N}} r^l(n) = \lim_{N\to\infty} \frac{1}{N} {}^t\mathbf{1}_N \mathbf{\Sigma}^{(1)} \mathbf{1}_N \geq \lim_{N\to\infty} \frac{C_*^{(1)}}{N} {}^t\mathbf{1}_N \mathbf{\Lambda}^l \mathbf{1}_N = C_*^{(1)} r^l(0).$$

From (20) and (46) it follows

$$\{2^{-m/2} r(m,n)\}^l \leq \{\Psi_{\gamma,H}(m) K_\gamma(m,n)\}^l \cdot \frac{1}{C_*^{(1)}} \sum_{n\in\mathbb{Z}} r_0^l(n).$$

Thus we have

$$R_H^l(m) = \sum_{n\in\mathbb{Z}} \rho^l(m,n) \leq \sum_{n\in\mathbb{Z}} [\Psi_{\gamma,H}(m) K_\gamma(m,n)]^l \cdot \frac{1}{C_*^{(1)}} R_H^{(1)}.$$

Lemma 10. *Let $\zeta = \zeta_1 > 0$ be defined as in* (45). *For $l \geq p$, $\zeta_l = \zeta_1^l$ is sufficient.*

REMARK 4. It is shown in [2, Lemma 3] that $r_0(n) \geq 0$.

8.2. Proof of Theorem 7. We have the values of $(C_{H,2}, C_{H,3}, C_{H,4})$ and (ξ_2, ξ_3, ξ_4) as

$$(C_{H,2}, C_{H,3}, C_{H,4}) = \left(1 - \frac{2^{-2H+1}}{3}, 1 - \frac{2^{-4H+1}}{5}, 1 + \frac{26}{87} \frac{2^{-2H}}{1 - 2^{-(2H+1)}}\right)$$

and $(\xi_2, \xi_3, \xi_4) = (3/2, 5/8, 87/200)$, respectively. In the first part of Proposition 6, we have

$$C_{H,J}\xi_J \leq \frac{v_{H,J}^2}{R_H} \leq \xi_J \quad \text{for} \quad J = 2, 3, \tag{47}$$

so that we have

$$\frac{v_{H,2}^2}{R_H} \geq C_{H,2}\xi_2 = \left(1 - \frac{2^{-2H+1}}{3}\right) \cdot \frac{3}{2} \geq \xi_3 = \frac{5}{8} \geq \frac{v_{H,3}^2}{R_H}.$$

Thus it turns out that $v_{H,2}^2 \geq v_{H,3}^2 \geq v_{H,4}^2$ for all $0 < H < 1$. By Proposition 6,

$$\begin{aligned}\frac{v_{H,3}^2}{R_H} &\geq C_{H,3}\xi_3 = \left(1 - \frac{2^{-4H+1}}{5}\right)\cdot\frac{5}{8}\\ &\geq C_{H,4}\xi_4 = \left(1 + \frac{26}{87}\frac{2^{-2H}}{1-2^{-(2H+1)}}\right)\cdot\frac{87}{200} \geq \frac{v_{H,4}^2}{R_H}.\end{aligned}$$

For $J \geq 4$, since

$$\xi_J \leq \frac{v_{H,J}^2}{R_H} \leq C_{H,J}\xi_J$$

(see Remark 1), the assertion is proved if

$$\xi_4 \geq C_{H,5}\xi_5 \tag{48}$$

and

$$C_{H,J}\xi_J \leq \xi_{J+1}, \quad \text{for all} \quad J = 5, 6, \ldots. \tag{49}$$

From the argument in the proof of Proposition 6, we evaluate $C_{H,J}$ in (15) through

$$\begin{aligned}&\sum_{m=1}^{\mathcal{J}-1}\frac{\xi_J(m)}{\xi_J}\cdot\frac{R_H(m)}{R_H}\\ &\leq \frac{1}{C_*^{(1)}}\sum_{m=1}^{\lfloor (J-2)/2\rfloor} 2^{-(2H+1)m}\cdot\Psi_{\gamma,H}(m)K_\gamma(m)\\ &\qquad\qquad \times\frac{2^J\{J^2-2(3+m)J+17+6m\}-2^m\{J^2+2(3-m)J+17-6m\}}{2^J(J^2-6J+17)-(J^2+6J+17)}.\end{aligned}$$

Taking $J = 5$ yields

$$\begin{aligned}C_{5,H} &= 1 + 2\cdot 2^{-(2H+1)}\cdot\frac{\xi_5(1)}{\xi_5}\cdot\frac{R_H(1)}{R_H}\\ &\leq 1 + 2^{-2H}\cdot\frac{36/400}{156/400}\cdot\frac{5/3}{C_*^{(2)}}[\Psi_{\gamma,H}(1)K_{\gamma,4}(1)]^2,\end{aligned}$$

which proves (48). A computation according to (34) yields $\sup_{0<H<1,\gamma\in\mathbb{N}} C_{5,H} = C_{5,1} \doteqdot 1.035$ for $\gamma = 1$, while $\xi_4 = 87/200 = 0.4350$ and $\xi_5 = 156/400 = 0.390$. Thus it turns out that every case satisfies (48). Similarly (49) with $J = 5$ is proven.

For $J = 6$, we have $\xi_6 = 999/2450 \fallingdotseq 0.4078$ and $\xi_7 = 1482/3136 \fallingdotseq 0.4726$, and

$$\begin{aligned} C_{6,H} &= 1+2\left\{2^{-(2H+1)}\frac{\xi_6(1)}{\xi_6}\cdot\frac{R_H(1)}{R_H}+2^{-2(2H+1)}\frac{\xi_6(2)}{\xi_6}\cdot\frac{R_H(2)}{R_H}\right\} \\ &= 1+2\cdot\frac{5/3}{C_*^{(2)}}\left\{2^{-(2H+1)}\frac{281}{999}[\Psi_{\gamma,H}(1)K_{\gamma,4}(1)]^2+2^{-2(2H+1)}\frac{27}{999}[\Psi_{\gamma,H}(2)K_{\gamma,4}(2)]^2\right\}. \end{aligned}$$

Hence $\sup_{0<H<1,\gamma\in\mathbb{N}} C_{6,\gamma,H} = C_{6,1} \fallingdotseq 1.051$ for $\gamma = 1$ and thus (49) with $J = 6$ indeed holds. The case $J = 7$ is treated in a similar way.

For $J \geq 8$, since $R_{\gamma,H}(m)$ is decreasing with respect to m, we can use a rough evaluation

$$C_{J,H} \leq 1 + 2\frac{R_H(1)}{R_H}\sum_{m=1}^{\lfloor (J-2)/2\rfloor} 2^{-(2H+1)m}\cdot\frac{L_J(m)}{L_J} \triangleq 1 + 2\frac{R_H(1)}{R_H}\cdot Z_{J,H}$$

to show (37) for $J = 8$. In fact, considering $Z_{J,H} - 2^{-(2H+1)}Z_{J,H}$, we obtain

$$Z_{J,H} \leq \frac{2^{-(2H+1)}}{1-2^{-(2H+1)}}\cdot\frac{2^J(J^2-6J+23)-2(J^2+4J+11)}{2^J(J^2-6J+17)-(J^2+6J+17)},$$

the right hand side of which is decreasing for $J \geq 7$. Hence

$$\sup_{J\geq 7} C_{J,H} \leq 1 + 2\frac{3698}{2964}\cdot\frac{2^{-(2H+1)}}{1-2^{-(2H+1)}}\cdot\frac{5/3}{C_*^{(2)}}[\Psi_{\gamma,H}(1)K_{\gamma,4}(1)]^2, \tag{50}$$

while

$$\min_{J\geq 7}\frac{\xi_{J+1}}{\xi_J} = \left(\frac{J-1}{J+2}\right)^2\frac{2^J(J^2-4J+12)-(J^2+8J+24)}{2^J(J^2-6J+17)-(J^2+6J+17)}\Bigg|_{J=7}, \tag{51}$$

the right hand side of which is equal to $455/741 \fallingdotseq 0.6140$. The evaluations (50) and (51) imply $C_{J,\gamma,H}\xi_J \leq \xi_{J+1}$ for $J \geq 7$, $\gamma \in \mathbb{N}$ and $0 < H < 1$. This completes the proof. □

REMARK 5. One cannot show (49) for the critical cases $J = 5$ and 6 by the rough evaluation based only on (50) and (51). The rough evaluation is however global for all $J \geq 5$, which is necessary for the global minimum.

9. Proof of Propositions

9.1. Proof of Proposition 1. The CLT itself has been proven in [2]. We prove (12) here. According to (14), the elements of $\Sigma = (\Sigma_{j,j'})$, $J_0 + 1 \leq j$, $j' = j + m \leq J$

is given as

$$
\begin{aligned}
\Sigma_{j,j'} &= \lim_{T\to\infty} \mathrm{Cov}\left[\sqrt{N_{j,T}}\frac{\hat{\theta}_{j,T}}{\theta_j}, \sqrt{N_{j',T}}\frac{\hat{\theta}_{j',T}}{\theta_j}\right] \\
&= \lim_{T\to\infty} \sqrt{\frac{N_{j',T}}{N_{j,T}}}\frac{1}{N_{j',T}} \sum_{k_1=1}^{N_{j,T}} \sum_{k_2=1}^{N_{j',T}} \mathrm{Cov}[H_l(s_0(k_1)), H_l(s_m(k_2))] \\
&= 2^{-m/2}\Bigg[\rho^l(m) + \lim_{T\to\infty}\frac{1}{N_{j',T}} \sum_{k=1}^{N_{j',T}-1} (N_{j',T} - \lfloor 2^{-m}k \rfloor)\rho^l(m,-k) \qquad (52) \\
&\qquad + \lim_{T\to\infty}\frac{1}{N_{j',T}} \sum_{k=1}^{N_{j',T}-1} (N_{j',T} - \lceil 2^{-m}k \rceil)\rho^l(m,k)\Bigg] \\
&= 2^{-m/2}\sum_{n\in\mathbb{Z}} \rho^l(m,n) = 2^{-m/2}R_H^{(1)}(m),
\end{aligned}
$$

where we have used a known result (31) and Lemma 8 above. (52) implies $\Sigma_{j,j+m} = \Sigma_{0,m}$. Consequently, $v_{H,J}^2 = \lim_{T\to\infty} \mathrm{Var}[\sqrt{T}\hat{\boldsymbol{\theta}}_T \cdot \dot{\boldsymbol{f}}]$ can be written as

$$
\begin{aligned}
v_{H,J}^2 &= \lim_{T\to\infty} \mathrm{Var}\left[\sqrt{T}\sum_{j\in\mathcal{J}} \hat{\theta}_{j,T}^{(1)} \cdot \dot{f}_j\right] = \lim_{T\to\infty} \mathrm{Var}\left[\sum_{j\in\mathcal{J}} \sqrt{N_{j,T}}\frac{\hat{\theta}_{j,T}^{(1)}}{\theta_j^{(1)}} \cdot 2^{j/2}\theta_j^{(1)}\dot{f}_j\right] \\
&= \sum_{j\in\mathcal{J}}\sum_{j'\in\mathcal{J}} \Sigma_{j,j'}2^{(j+j')/2}a_j a_{j'}[\theta_j\dot{\varphi}_j(\theta_j)][\theta_{j'}\dot{\varphi}_{j'}(\theta_{j'})],
\end{aligned}
$$

which yields the desired equation. □

9.2. Proof of Proposition 4. For $m = 1$, we have

$$
K_\gamma^2(1,n) = \sum_{k\in\mathbb{N}} \frac{1}{\{1+(n+\delta k)^2\}^{2\gamma}} \le \frac{1}{\{1+(n+\delta)^2\}^{2\gamma}} + \int_1^\infty \frac{dx}{\{1+(n+\delta x)^2\}^{2\gamma}}.
$$

Although $\delta = 2$ in the case of $m = 1$, we keep the variable δ below, since the calculation here will be reused for the case of $m = 2$ (i.e. $\delta = 4$) as well.

Since the last integral can be evaluated as

$$
\begin{aligned}
\int_1^\infty \frac{dx}{\{1+(n+\delta x)^2\}^{2\gamma}} &= \frac{1}{\delta}\int_{n+\delta}^\infty \frac{dx}{(1+x^2)^{2\gamma}} = \frac{1}{\delta}\int_{n+\delta}^\infty \frac{(1+x^2)^{1/2}}{2x} \cdot \frac{2x}{(1+x^2)^{2\gamma+1/2}}\,dx \\
&\le \frac{1}{\delta(4\gamma-1)(n+\delta)} \cdot \frac{1}{\{1+(n+\delta)^2\}^{2\gamma-1}},
\end{aligned}
$$

we have

$$
K_\gamma^2(1,n) \le \frac{\bar{c}_{\gamma,\delta}^{(K1)}}{(n+\delta)\{1+(n+\delta)^2\}^{2\gamma-1}}, \quad \bar{c}_{\gamma,\delta}^{(K1)} = \left[\frac{1}{\delta(4\gamma-1)} + \frac{\delta}{1+\delta^2}\right],
$$

which leads, taking $n=0$ and $\gamma=1$ in $\bar{c}^{(K1)}_{\gamma,\delta}$ as well as taking $\delta=2$, to (25).

Therefore, for $K^l_\gamma(1)=K^l_\gamma(1,0)+2\sum_{n\in\mathbb{N}}K^l_\gamma(1,n)$, we have

$$K^l_\gamma(1)\le[\bar{c}^{(K1)}_{\gamma,\delta}]^{l/2}\left[\left(\frac{1}{\delta}\right)^{l/2}\left(\frac{1}{1+\delta^2}\right)^q+2\sum_{n\in\mathbb{N}}\frac{1}{(n+\delta)^{l/2}\{1+(n+\delta)^2\}^q}\right]. \tag{53}$$

The last summation can be evaluated as

$$\begin{aligned}&\sum_{n\in\mathbb{N}}\frac{1}{(n+\delta)^{l/2}[1+(n+\delta)^2]^q}\\&\le\int_\delta^\infty\frac{dx}{x^{l/2}(1+x^2)^q}=\int_\delta^\infty\frac{(1+x^2)^{(l+2)/4}}{x^{l/2}\cdot 2x}\cdot\frac{2x}{(1+x^2)^{q+(l+2)/4}}\,dx\\&\le\frac{4}{2(4q+l-2)}\cdot\left(\frac{1}{\delta}\right)^{(l/2)+1}\left(\frac{1}{1+\delta^2}\right)^{q-1},\end{aligned}$$

and, applying this to (53), we obtain $K^l_\gamma(1)\le c^{(K1)}_{\gamma,\delta}[(3/5)\cdot(1/\delta)(1/(1+\delta^2))^{2\gamma-1}]^{l/2}$. Taking $\delta=2$, we get (26). For $m=2$, we have

$$\begin{aligned}K^2_\gamma(2,n)&=\sum_{k_1,k_2\in\mathbb{N}}\frac{1}{[1+(n+4k_1+4k_2)^2]^{2\gamma}}=\sum_{k\in\mathbb{N}}\frac{k-1}{[1+(n+4k)^2]^{2\gamma}}\\&\le\int_1^\infty\frac{x-1}{[1+(n+4x)^2]^{2\gamma}}\,dx=\frac{1}{16}\int_{n+4}^\infty\frac{x}{(1+x^2)^{2\gamma}}\,dx-\frac{1}{4}\int_{n+4}^\infty\frac{dx}{(1+x^2)^{2\gamma}}.\end{aligned}$$

Ignoring the second integral term in the right hand side, we have (27). The evaluation of $\bar{K}^2_\gamma(2,n)$ is given by taking $\delta=4$ in (53). Finally, the evaluation in (28) for $K^2_\gamma(2)\le(1/17)^q+2\sum_{n\in\mathbb{N}}[1+(n+4)^2]^{-q}$ follows from

$$\sum_{n\in\mathbb{N}}[1+(n+4)^2]^{-q}\le\int_4^\infty\frac{dx}{(1+x^2)^q}=\int_4^\infty\frac{(1+x^2)^{1/2}}{2x}\cdot\frac{2x}{(1+x^2)^{q+1/2}}\,dx,$$

the right hand side of which is bounded by $1/(2(4q-2))(1/17)^q$. □

9.3. Proof of Proposition 5. Recall that

$$|m_0(\lambda)|^2=\left(\cos^2\frac{\lambda}{2}\right)^\gamma P_\gamma\left(\sin^2\frac{\lambda}{2}\right),\quad\text{where}\quad P_\gamma(y)=\sum_{n=0}^{\gamma-1}\binom{\gamma-1+n}{n}y^n,$$

for an MRA wavelet of γ-th order. We remark that $P_1(y)\equiv 1$ and the relation $|m_0(\lambda)|^2+|m_0(\lambda+\pi)|^2\equiv 1$, or what is the same,

$$\left(\cos^2\frac{\lambda}{2}\right)^\gamma P_\gamma\left(\sin^2\frac{\lambda}{2}\right)+\left(\sin^2\frac{\lambda}{2}\right)^\gamma P_\gamma\left(\cos^2\frac{\lambda}{2}\right)\equiv 1 \tag{54}$$

holds (see [8]). For (29), we have

$$\sum_{k\in\mathbb{N}} |\alpha^{(1)}_{\gamma,k}|^2 = \frac{1}{\pi}\int_{2\pi}^{3\pi} |\varphi^{(1)}_\gamma(\lambda)|^2\, d\lambda = \frac{1}{\pi}\int_{2\pi}^{3\pi} \frac{|m_0(\lambda+\pi)|^2 |m_0|(\lambda/2)|^2}{|m_0(\lambda/2+\pi)|^2}\, d\lambda$$

$$= \frac{1}{\pi}\int_0^{\pi} \frac{(2\sin^2\lambda/4)^{2\gamma} P_\gamma(\cos^2\lambda/2) P_\gamma(\cos^2\lambda/4)}{P_\gamma(\sin^2\lambda/4)}\, d\lambda \triangleq \frac{1}{\pi}\int_0^{\pi} g^{(1)}_\gamma(\lambda)\, d\lambda,$$

by Parseval's equality (see (40)). Similarly, for (30), we have

$$\sum_{k\in\mathbb{N}} |\alpha^{(2)}_{\gamma,k}|^2 = \frac{1}{\pi}\int_0^{\pi} |\varphi^{(2)}_\gamma(\lambda)|^2\, d\lambda = \frac{1}{\pi}\int_0^{\pi} \frac{|m_0(\lambda/2+\pi)|^2}{|m_0(\lambda+\pi)|^2 |m_0(\lambda/2)|^2}\, d\lambda$$

$$= \frac{1}{\pi}\int_0^{\pi} \frac{P_\gamma(\cos^2\lambda/4)}{(2\cos^2\lambda/4)^{2\gamma} P_\gamma(\cos^2\lambda/2) P_\gamma(\sin^2\lambda/4)}\, d\lambda \triangleq \frac{1}{\pi}\int_0^{\pi} g^{(2)}_\gamma(\lambda)\, d\lambda.$$

We will show that the functions $g^{(i)}_\gamma(\lambda)$, $i = 1, 2$ are convex downward. If this is shown, upper bounds of the two, by piecewise linear segments, are valid. We can especially take the linear segments over $[0, 2\pi/3]$ and $[2\pi/3, \pi]$ (see Remark 7 below). Then we have, by the area formula for a trapezoid,

$$\begin{aligned} &\frac{1}{\pi}\int_0^{\pi} g^{(i)}_\gamma(\lambda)\, d\lambda \\ &\le \frac{1}{\pi}\left[\frac{1}{2}\cdot\frac{2\pi}{3}\cdot\left\{g^{(i)}_\gamma(0) + g^{(i)}_\gamma\frac{2\pi}{3}\right\} + \frac{1}{2}\cdot\frac{\pi}{3}\cdot\left\{g^{(i)}_\gamma\frac{2\pi}{3} + g^{(i)}_\gamma(\pi)\right\}\right], \end{aligned} \tag{55}$$

for $i = 1, 2$. By inspection, we have $g^{(1)}_\gamma(0) = 0$, $g^{(1)}_\gamma(\pi) = 1 = g^{(2)}_\gamma(\pi)$ and $g^{(2)}_\gamma(0) = 4^{-\gamma}$. Using these values, we obtain (29) and (30).

It remains to show that the convexity of $g^{(1)}_\gamma(\lambda)$ and $g^{(2)}_\gamma(\lambda)$ holds. For $g^{(1)}_\gamma(\lambda)$, this follows from the convexity of the two terms $(\sin^2\lambda/4)^\gamma \cdot P_\gamma(\cos^2\lambda/4)$ and $(\sin^2\lambda/4)^\gamma \cdot P_\gamma(\cos^2\lambda/2)/P_\gamma(\sin^2\lambda/4)$. In fact,

$$\left(\sin^2\frac{\lambda}{4}\right)^\gamma P_\gamma\left(\cos^2\frac{\lambda}{4}\right) = \sum_{\nu=0}^{\gamma-1}\binom{\gamma-1+\nu}{\nu}\left(\frac{1}{4}\sin^2\frac{\lambda}{2}\right)^\nu\cdot\left(\sin^2\frac{\lambda}{4}\right)^{\gamma-\nu},$$

which is convex downward on $[0, \pi]$, since each summand is so. Also,

$$\begin{aligned} \frac{(2\sin^2\lambda/4)^\gamma P_\gamma(\cos^2\lambda/2)}{P_\gamma(\sin^2\lambda/4)} &= \frac{(2\sin^2\lambda/4)^\gamma\cdot(\sin^2\lambda/2)^\gamma P_\gamma(\cos^2\lambda/2)}{(\sin^2\lambda/2)^\gamma P_\gamma(\sin^2\lambda/4)} \\ &= \frac{(\sin^2\lambda/2)^\gamma P_\gamma(\cos^2\lambda/2)}{(2\cos^2\lambda/4)^\gamma P_\gamma(\sin^2\lambda/4)} \\ &= \frac{(\sin^2\lambda/2)^\gamma P_\gamma(\cos^2\lambda/2)}{4^\gamma[1-(\sin^2\lambda/4)^\gamma P_\gamma(\cos^2\lambda/4)]}, \end{aligned}$$

where we have used (54). In the last fraction, the numerator is convex downward and the denominator is convex upward so that the fraction is convex downward.

Similarly, the convexity of $g_\gamma^{(2)}$ follows from

$$g_\gamma^{(2)}(\lambda) = \frac{(\cos^2\lambda/4)^{-\gamma} P_\gamma(\cos^2\lambda/4)}{4^\gamma P_\gamma(\cos^2\lambda/2)[1-(\sin^2\lambda/4)^\gamma P_\gamma(\cos^2\lambda/4)]},$$

which itself is obtained by multiplying $(\sin^2\lambda/4)^\gamma$ the numerator and denominator of $g_\gamma^{(2)}$ and using (54). □

REMARK 6. The values of $|\varphi_\gamma^{(1)}(\lambda)|^2$ at $\lambda = 0$ and $|\varphi_\gamma^{(2)}(\lambda)|^2$ at $\lambda = 2\pi$ are large: $\lim_{\lambda\to 2\pi}|\varphi_\gamma^{(1)}(\lambda)|^2 = 4^\gamma$ and $\lim_{\lambda\to 2\pi}|\varphi_\gamma^{(2)}(\lambda)|^2 = \infty$. The large values cannot be used in the evaluation like (55). Therefore we consider $\varphi_\gamma^{(1)}(\lambda)$ on $[\pi, 3\pi]$ and $\varphi_\gamma^{(2)}(\lambda)$ on $[-\pi, \pi]$ respectively, where they are bounded by 1.

REMARK 7. The special value $\lambda = 2\pi/3$ which appeared in the proof of Proposition 5 is related to the invariant cycles ([8, p. 188]) for the mapping $\tau : \tau\lambda = 2\lambda \pmod{2\pi}$. Here we have used $\lambda = 2\pi/3$ in order to make the calculations easier in the expressions in which the arguments $\lambda/4$ and $\lambda/2$ are involved.

In the upper bound of the integral on $[0, \pi]$ in (55), we have taken the linear segments over $[0, 2\pi/3]$ and $[2\pi/3, \pi]$. Whether or not there exists a more convenient and precise way of segmentation, in which a trade-off between convenience and precision may be involved, is to be clarified.

9.4. Proof of Proposition 6. For $J = 2$ or 3, since $\xi_2(1)/\xi_2 = -1/3$ and $\xi_3(1) = 0$, $\xi_3(2)/\xi_3 = -1/5$, we have

$$v_{H,J}^2 = \xi_J R_H^2\left[1 + 2\sum_{m=1}^{J-1}\frac{2^{-2Hm}\xi_J(m)}{\xi_J}\cdot\frac{\bar R_H^2(m)}{\bar R_H^2}\right] \ge C_{H,J}\xi_J R_H^2$$

by (47) and Lemma 11 below, where $C_{H,2} = 1 - 2^{1-2H}/3$ and $C_{H,3} = 1 - 2^{1-4H}/5$. This proves (36).

To show that the lower bound in (37) holds, we check the positivity of $U_H(J) \triangleq \sum_{m=1}^{J-1} 2^{-2Hm}\xi_J(m)\bar R_H^2(m)$ for $J \ge 4$. If this is checked, then $v_{H,J}^2$ is estimated from below, by ignoring $U_H(J)$, only by the diagonal part $\xi_J R_H^2$. To this end we note that $U_H(J)$, $J \ge 4$ satisfies the recurrence relation

$$\begin{aligned} U_H(J) &= 2U_H(J-2) + 2^{-2H(J-1)}a_J^2(1)\bar R_H^2(J-1) \\ &\quad + \sum_{m=1}^{J-1} 2^{-2Hm}(2+2^{J-m})a_J(1)a_J(1+m)\bar R_H^2(m). \end{aligned} \tag{56}$$

By inspection, using (34) and (35), we have that $U_H(4)$ is bounded below by

$$2^{-2H}\frac{13}{2}\bar{R}_H^2(1) + 2^{-4H}\frac{9}{2}\bar{R}_H^2(2) - 2^{-6H}\frac{9}{2}\bar{R}_H^2(3) \geq 2^{-2H}\frac{25}{8}\bar{R}_H^2(1) > 0$$

and similarly $U_H(5) \geq 2^{-2H}30\bar{R}_H^2(1) > 0$. Also the sum on the right hand side of (56) is bounded below by

$$\begin{aligned}&(2\bar{R}_H^2(1) + 2^{J+1}\bar{R}_H^2(J+1))\left\{\frac{(J+1)^2}{2}\sum_{m=1}^{J+1} 2^{-(2H+1)m} - (J+1)\sum_{m=1}^{J+1} 2^{-(2H+1)m} m\right\}\\ &\geq (\bar{R}_H^2(1) + 2^J\bar{R}_H^2(J+1))(J+1)(J-3),\end{aligned}$$

which is positive for $J \geq 6$. Hence, from (56), the positivity of $U_H(J)$ for even J $(J \geq 6)$ and odd J $(J \geq 7)$ follows, respectively.

To prove the upper bound in (37), we remark that it follows, by neglecting the negative terms,

$$\frac{U_H(J)}{\xi_J \bar{R}_H^2} \leq \frac{\sum_{m=1}^{\lfloor (J-2)/2\rfloor} \xi_J(m)}{\xi_J} \leq \frac{\sum_{m=1}^{\lfloor (J-2)/2\rfloor} 2^{-2Hm} L_J(m)}{L_J(0)}, \tag{57}$$

where $L_J(m) \triangleq \sum_{j=1}^{J-m} 2^j (x_j - \bar{x}_J)(x_{j+m} - \bar{x}_J)$ is calculated to be

$$\sum_{j=1}^{J-m} 2^j\left\{j^2 - (J+1-m)j + \frac{(J+1)(J+1-m)}{4}\right\}, \quad m \in \mathbb{N}_0,$$

and where we have used only the positive terms and the fact that $R_H^2(m)$ is decreasing with respect to m in (57). By Lemma 11 below, the right hand side of (57) is equal to

$$\sum_{m=1}^{\lfloor (J-2)/2\rfloor} 2^{-(2H+1)m} \cdot \frac{2^J\{J^2-2(3+m)J+17+6m\}-2^m\{J^2+2(3-m)J+17-6m\}}{2^J(J^2-6J+17)-(J^2+6J+17)}.$$

Therefore, taking $C_{H,J} = 1 + 2[U_H(J)/(\xi_J\bar{R}_H^2)]$ and applying the dominated convergence theorem as $J \to \infty$ for the summation yields (38). $\square$

Lemma 11. *For $m \in \mathbb{N}_0$ and $J \in \mathbb{N}_0$, $J \geq 2$, we have*

$$L_J(m) = 2^{J-m-1}\{J^2 - 2(3+m)J + 17 + 6m\} - 2^{-1}\{J^2 + 2(3-m)J + 17 - 6m\}$$

and

$$\xi_J \leq \frac{18[(2^J(J^2 - 6J + 17) - (J^2 + 6J + 17)]}{[(J-1)J(J+1)]^2}.$$

10. Proof of Lemmas

10.1. Proof of Lemma 8. Since

$$2^{-m/2}r(m,n) = C_H \int_{(0,\infty)} e^{in\lambda}\hat{\psi}(\lambda)\overline{\hat{\psi}(2^m\lambda)}\,d\mu_H(\lambda)$$

is real-valued, we have

$$\begin{aligned}
&2^{-m/2}r(m,n)\\
&= 2^{-m/2}\frac{r(m,n)+\overline{r(m,n)}}{2}\\
&= \frac{C_H}{2}\left[\int_{(0,\infty)} e^{in\lambda}\hat{\psi}(\lambda)\overline{\hat{\psi}(2^m\lambda)}\,d\mu_H(\lambda) + \int_{(0,\infty)} e^{-in\lambda}\overline{\hat{\psi}(\lambda)}\hat{\psi}(2^m\lambda)\,d\mu_H(\lambda)\right]\\
&= \frac{C_H}{2}\int_{\mathbb{R}\setminus 0} e^{in\lambda}\hat{\psi}(\lambda)\overline{\hat{\psi}(2^m\lambda)}\,d\mu_H(\lambda),
\end{aligned}$$

where we have redefined μ_H by $d\mu_H(\lambda) = |\lambda|^{-2H}\,d\lambda$ on $\mathbb{R}\setminus 0$. Similarly, starting the same process by writing $2^{-m/2}r(m,n) = C_H \int_{(0,\infty)} e^{in\lambda}\overline{\hat{\psi}(\lambda)}\hat{\psi}(2^m\lambda)\,d\mu_H(\lambda)$, we have

$$2^{-m/2}r(m,n) = \frac{C_H}{2}\int_{\mathbb{R}\setminus 0} e^{in\lambda}\overline{\hat{\psi}(\lambda)}\hat{\psi}(2^m\lambda)\,d\mu_H(\lambda).$$

Adding the two displayed identities, we obtain

$$2^{-m/2}r(m,n) = (C_H/2)\int_{\mathbb{R}\setminus 0} e^{in\lambda}\,\mathfrak{Re}[\hat{\psi}(\lambda)\overline{\hat{\psi}(2^m\lambda)}]\,d\mu_H(\lambda).$$

Thus we can replace $e^{in\lambda}$ by $\cos(n\lambda)$, so that $r(m,-n) = r(m,n)$. □

10.2. Proof of Lemma 9. Recall the expression (3) of m_0. Let us consider the 4π-periodic function defined by $\overline{\hat{\psi}(2\lambda)}\big/\overline{\hat{\psi}(\lambda)}$, $\lambda \in [2\pi, 3\pi]$; $= 0$, $\lambda \in [0, 2\pi) \cup (3\pi, 4\pi]$. Substituting (3), we have formally, on $[2\pi, 3\pi]$,

$$\begin{aligned}
\frac{\overline{\hat{\psi}(2\lambda)}}{\overline{\hat{\psi}(\lambda)}} &= e^{-i(N_2-N_1-1)\lambda/2}\cdot\frac{\sum_{k=0}^{N_2-N_1}[h_{N_2-k}(-1)^k/h_{N_2}]e^{ik\lambda}\cdot\sum_{k=0}^{N_2-N_1}\bar{h}_{k+N_1}e^{ik\lambda/2}}{1+\sum_{k=1}^{N_2-N_1}[h_{N_2-k}(-1)^k/h_{N_2}]e^{ik\lambda/2}}\\
&= e^{-i(N_2-N_1-1)\lambda/2}\varphi^{(1)}(\lambda),\\
\varphi^{(1)}(\lambda) &= \sum_{k\in\mathbb{N}}\tilde{\alpha}_k^{(1)}e^{ik\lambda/2}
\end{aligned}$$

for some $\{\tilde{\alpha}_k^{(1)};\, k \in \mathbb{N}\}$. The Fourier coefficient of the 4π-periodic function $\varphi^{(1)}(\lambda)$ is, among such $\{\tilde{\alpha}_k^{(1)}\}$, the one that makes the Fourier series $\sum_{k\in\mathbb{N}}\tilde{\alpha}_k^{(1)}e^{ik\lambda/2}$ vanish on

$[0,2\pi)\cup(3\pi,4\pi]$. Let us denote such a Fourier coefficient by the same symbol $\{\tilde{\alpha}_k^{(1)}\}$. Then, the Fourier coefficient $\{\tilde{\alpha}_k^{(1)}\}$ of $\varphi^{(1)}(\lambda)$ on $[0,4\pi]$ is given by

$$\tilde{\alpha}_k^{(1)} = \frac{1}{4\pi}\int_0^{4\pi} \varphi^{(1)}(\lambda)e^{-ik\lambda/2}d\lambda = \frac{1}{4\pi}\int_{2\pi}^{3\pi} \varphi^{(1)}(\lambda)e^{-ik\lambda/2}d\lambda. \tag{58}$$

Thus we have the Fourier series of $\varphi_1^{(1)}(\lambda)$ on $[0,4\pi]$, $\varphi^{(1,2)}(\lambda) = \sum_{k\in\mathbb{N}} \tilde{\alpha}_k^{(1)} e^{ik\lambda/2}$.

On the other hand, $\varphi^{(1)}(\lambda)$ on $[2\pi,3\pi]$ can be written as a Fourier series on an interval of π-length, $\varphi^{(1)}(\lambda) = \sum_k \alpha_k^{(1)} e^{i2k\lambda}$, $\lambda\in[2\pi,3\pi]$ for some coefficient $\{\alpha_k^{(1)}\}$. This $\{\alpha_k^{(1)}\}$ is given by $\alpha_k^{(1)} = (1/\pi)\int_{2\pi}^{3\pi}\varphi^{(1)}e^{-i2k\lambda}d\lambda = 4\tilde{\alpha}_{4k}^{(1)}$, $k\in\mathbb{Z}$ by (58). Therefore $\alpha_k^{(1)} \equiv 0$ for $k\in-\mathbb{N}_0$. □

10.3. Proof of Lemma 10. Recall that ζ_l and C_{l*} are the constants in (45). We evaluate ζ_l, $l\in\mathbb{N}$ using $\zeta_1>0$. We have $\sum_{k\in\mathbb{N}}(\rho_k^2)^l \le \left[\sum_{k\in\mathbb{N}}\rho_k^2\right]^l \le (1/\zeta_1^2)^l$. Thus, as an upper bound for $1/\zeta_l^2$ we can take $1/\zeta_1^{2l}$, i.e. $\zeta_l = \zeta_1^l$ is sufficient. Hence $C_*^{(l)} = (\zeta_l-1)/\zeta_l = (\zeta_1^l-1)/\zeta_1^l$. □

10.4. Proof of Lemma 11. The equality for $L_J(m)$ is obtained by making $2L_J(m)-L_J(m)$. Since $\sum_{j=1}^J(x_j-\bar{x}_J)^2 = (J-1)J(J+1)/12$ and $L_J(m)$ is decreasing with respect to $m\in\mathbb{N}_0$, we also have $\xi_J \le L_J(0)\big/\left[2\sum_{j=1}^J(x_j-\bar{x}_J)^2\right]^2$, which yields the inequality for ξ_J. □

11. Concluding remarks

In this paper, we have formulated a j-localization property of wavelet coefficients of FBM. So far in the relevant field, only k-localization have been considered essentially and how to formulate the j-localization have been left unsolved. Our contribution here is three fold: First, in the basic j-localization theorem, we evaluated "pointwise" the cross-scale covariance of WC and evaluated two key elements $K_{\gamma,H}(m,n)$ and $\Psi_{\gamma,H}(m)$ in the covariance.

Second, we formulated, as a typical example of the functional form of j-localization, the limiting variance in the CLT of WCD estimates. It involves the evaluation of $R_H(m) = \sum_{k\in\mathbb{Z}}\rho^l(m,k)$. Because R_H, $m=1,2,\ldots$ are desired to decrease fast and indeed small, one has to obtain the "pointwise" evaluation of $\rho(m,k)$; the asymptotic evaluation as $k\to\infty$ considered by many authors so far does not work.

Third, as an application of the functional form of the j-localization, we found the best upper bound J of the scales $j=1,2,\ldots,J$ used in the Hurst index estimates, that makes the estimation variance minimum.

One of the important merits of wavelet method for statistical estimation is undoubtedly in the time-frequency localization. Original process X has argument of time, while

WC scale j and shift k. We can obtain the localization in price of increasing the original single argument, t, to the two, scale and shift.

ACKNOWLEDGMENT. The authors are grateful to the reviewer for valuable comments. Also, the second author would like to express his heartfelt appreciation to Professor Nakahiro Yoshida at University of Tokyo for encouragement and suggestions.

References

[1] P. Abry and D. Veitch: *Wavelet analysis of long-range-dependent traffic*, IEEE Trans. Inform. Theory **44** (1998), 2–15.

[2] S. Albeverio and S. Kawasaki: *On a localization property of wavelet coefficients for processes with stationary increments, and applications*, I, *Localization with respect to shift*, Adv. in Appl. Probab. **37** (2005), 938–962.

[3] J.M. Bardet, G. Lang, E. Moulines and P. Soulier: *Wavelet estimator of long-range dependent processes*, Stat. Inference Stoch. Process. **3** (2000), 85–99.

[4] J. Beran: Statistics for Long-Memory Processes, Chapman and Hall, New York, 1994.

[5] P. Billingsley: Probability and Measure, third edition, New York, 1995.

[6] P. Billingsley: Convergence of Probability Measures, second edition, Probability and Statistics, Wiley, New York, 1999.

[7] C. Chiann and P.A. Morettin: *A wavelet analysis for time series*, J. Nonparametr. Statist. **10** (1999), 1–46.

[8] I. Daubechies: Ten Lectures on Wavelets, third printing, SIAM, Philadelphia, PA, 1994.

[9] P.L. Duren: Theory of H^p Spaces, Dover Publications, 2000.

[10] P. Flandrin: *Wavelet analysis and synthesis of fractional Brownian motion*, IEEE Trans. Inform. Theory **38** (1992), 910–917.

[11] U. Grenander and M. Rosenblatt: Statistical Analysis of Stationary Time Series, second edition, Chelsea, New York, 1984.

[12] W. Härdle, G. Kerkyacharian, D. Picard and A. Tsybakov: Wavelets, Approximation, and Statistical Applications, Lecture Notes in Statistics **129**, Springer, New York, 1998.

[13] S. Kawasaki: $1/f$*-noise in communication networks—self-similar processes and long-range dependent processes—and their wavelet analysis*, Bull. Soc. Instr. Control Eng. **49** (2010), 236–243, (in Japanese).

[14] E.L. Lehmann: Elements of Large-Sample Theory, Springer, New York, 1999.

[15] E. Masry: *The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion*, IEEE Trans. Inform. Theory **39** (1993), 260–264.

[16] P. Abry, P. Flandrin, M.S. Taqqu and D. Veitch: *Self-similarity and long-range dependence through the wavelet lens*; in Theory and Applications of Long-Range Dependence, Birkhäuser, Boston, Boston, MA, 527–556, 2002.

[17] K. Tanaka: *Wavelet analysis*; in Statistics of Economical Time Series–Mathematical Foundation, Iwanami Publ., Tokyo, 2003, (in Japanese).

[18] K. Tanaka: *Applications of wavelet analysis to statistics*, Sūgaku **57** (2005), 50–69.

[19] A.H. Tewfik and M. Kim: *Correlations structure of the discrete wavelet coefficients of fractional Brownian motion*, IEEE Trans. Inform. Theory **38** (1992), 904–909.

[20] Z. Xie and T. Suzuki: *Wavelet and stochastic processes*, Uchida-Rokakuho Publ., Tokyo, 2002, (in Japanese).

Sergio Albeverio
Institut für Angewandte Mathematik
Universität Bonn
D-53115 Bonn
HCM, IZKUS, SFB611; BiBos; CERFIM (Locarno);
ENDENICHER ALLE 60
Germany
e-mail: albeverio@iam.uni-bonn.de

Shuji Kawasaki
Faculty of Humanities and Social Sciences
Iwate University
Morioka 020-8550
Japan
e-mail: shuji@iwate-u.ac.jp